SCIENCE OF LIGHT

VOLUME 9 NUMBER 1

June
1960

Published by the

Institute for Optical Research

Tokyo University of Education

in collaboration with

The Spectroscopical Society of Japan

SCIENCE OF LIGHT

Science of Light contains reports of the Institute for Optical Research and contribution from other science bodies about similar subjects.

All communications should be addressed to the director or to the librarian of the Institute.

The Institute for Optical Research

Tokyo University of Education

400, Hyakunin-tyo-4, Shinzyuku-ku, Tokyo, Japan

Printed by
Kabushiki Kaisha Kokusai Insatsu,
Tokyo

Plane Glating Monochromators of Ebert, Pfund and Czerny-Turner types

Keiei KUDO

Institute for Optical Research, Tokyo University of Education, Shinjuku-ku, Tokyo

(Received June 21, 1960)

Abstract

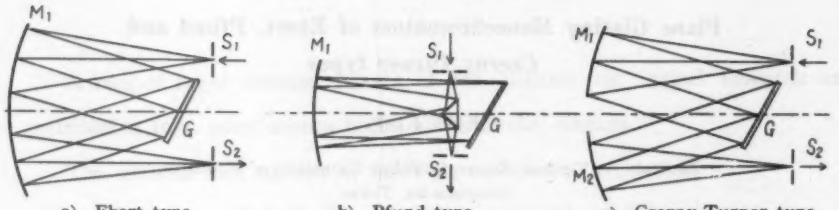
The aberration of plane grating monochromators such as of Ebert, Czerny-Turner and Pfund types are calculated in the case of symmetric arrangement. By using appropriate focal condition and grating equation, the slit form for avoiding the wave-length error is determined. It is a circular arc in Ebert and Pfund types and is an elliptic arc in non-concentric Czerny-Turner type. The length of the slit restricted by astigmatism and the optimum size of the grating derived from spherical aberration are determined.

1. Introduction

According to William G. Fastie,[1] both the entrance and exit slits in Ebert system are to be curved in order to remove the effect of astigmatism and to avoid wavelength error at the exit slit caused by the curvature of spectral lines when a plane grating is used. He showed that the resolving power obtained with any short portion of a long slit properly curved was the same as that with any other portion of the slit. In infrared region, in which radiation energy from the source is weak and detector is not sensitive, slits of monochromator must be long in order to let much light energy enter. Therefore, for an infrared spectrometer, long slits are very desirable. In recent years Ebert type spectrometers with long circular-arc-shaped slits have been constructed by many workers[2] and excellent results were reported. But it is surprizing that the aberration of Ebert type monochromator has not been seriously discussed, and that information about grating equation, focal condition, astigmatism and optimum grating size is still lacking.

Since Ebert or Pfund type monochromator (Fig. 1) incorporating with one concave mirror as collimator is a special case of Czerny-Turner type which has two concave mirrors, we shall consider here only the two mirror system. In such a system, be-

1) W. G. Fastie: J. Opt. Soc. Am. **42**, 641 and 647 (1952).
2) For example; M. A. Ford, W. C. Price and G. R. Wilkinson: J. Sci. Inst. **35**, 55 (1957). D. W. Robinson: J. Opt. Soc. Am. **49**, 966 (1959).



a) Ebert type b) Pfund type c) Czerny-Turner type

Fig. 1. Plane grating monochromator using spherical mirrors as collimators.

cause of numbers of adjustments being involved, the two spherical mirrors are apt to be so placed that their centers of curvature are not coincident. For such non-concentric arrangement, designers of monochromators should be prepared to face the problems whether the curved slit is applicable, how the grating equation and focal condition are affected and if the astigmatism is enhanced or the higher order aberration can vanish. To enlighten them on these subjects, it is necessary to examine the aberration in two mirror system in general. For this, application of Beutler's theory[3] on concave grating to the case of plane grating monochromator will serve.

2. General Expression

As shown in Fig. 2, the origin O of a rectangular coordinate system is taken at an arbitrary point on a holizontal line (x-axis) which passes through the grating center, and z-axis is in the vertical plane and parallel to the grating rullings. Light from a point $(x_1\, y_1\, z_1)$ on the entrance slit S_1 falls on the point $X(x_2\, y_2\, z_2)$ of the surface of mirror M_1 and reflected to the point $G(\xi\, \eta\, l)$ on the plane grating and next diffracted to the point $X'(x_2'\, y_2'\, z_2')$ on the mirror M_2 and finally reflected to the point $(x_1'\, y_1'\, z_1')$ on the image plane. Radii of curvatures of the mirrors M_1 and M_2 are R and R' respectively, centers of the curvature being at $O_1(x_0\, y_0\, O)$ and $O_2(x_0'\, y_0'\, O)$ respectively. Grating center is at $'COO)$, its width w and the angle between its surface and y-axis θ.

Now let the slit S_1, the mirror M_1 and the grating compose M_1-system and the grating, the mirror M_2 and the image plane, M_2-system. If we take the point O_1 as the origin for M_1-system by parallel transformation of the coordinates, we have for the points on slit S_1, mirror M_1 and grating G

$$\left.\begin{array}{lllll} \text{for slit,} & s_1 \equiv x_1 - x_0\,, & s_2 \equiv y_1 - y_0\,, & z_1 = z_1 & \\ \text{for mirror} & x \equiv x_2 - x_0\,, & y \equiv y_2 - y_0\,, & z \equiv z_2 & \\ \text{for grating} & G_1 \equiv \xi - x_0 = h_0 - w\phi_1\,, & G_2 \equiv \eta - y_0 = w\phi_2 - y_0\,, & l = l & \end{array}\right\} \quad (1)$$

3) H.G. Beutler: J. Opt. Soc. Am. **35**, 113 (1945).

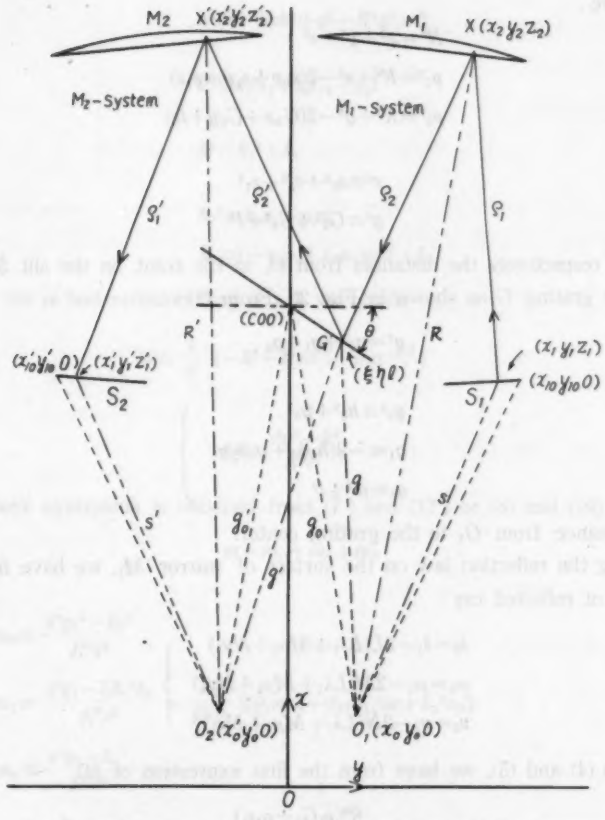


Fig. 2. Coordinates of optical elements in two mirrors system.

where,

$$h_0 \equiv c - x_0 , \qquad \phi_1 \equiv \sin\theta , \qquad \phi_2 \equiv \cos\theta$$

Then the direction consines of radius R of mirror M_1, incident light path ρ_1 and reflected light path ρ_2 become

$$\left.\begin{array}{lll} L=\dfrac{x}{R} & M=\dfrac{y}{R} & N=\dfrac{z}{R} \\ \lambda_1=\dfrac{x-s_1}{\rho_1} & \mu_1=\dfrac{y-s_2}{\rho_1} & \nu_1=\dfrac{z-z_1}{\rho_1} \\ \lambda_2=\dfrac{G_1-x}{\rho_2} & \mu_2=\dfrac{G_2-y}{\rho_2} & \nu_2=\dfrac{l-z}{\rho_2} \end{array}\right\} \qquad (2)$$

and therefore

$$R^2=x^2+y^2+z^2 \tag{3}$$

$$\rho_1^2=R^2+s^2-2(s_1x+s_2y+z_1z) \tag{4}$$

$$\rho_2^2=R^2+g^2-2(G_1x+G_2y+lz) \tag{5}$$

where,

$$s^2\equiv s_1^2+s_2^2+z_1^2 \tag{6}$$

$$g^2\equiv G_1^2+G_2^2+l^2 \tag{7}$$

s and g are respectively the distances from O_1 to the point on the slit S_1 and to the point on the grating G as shown in Fig. 2. From (1)

$$g^2=g_0^2+g_1+g_2 \tag{7'}$$

where,

$$\left.\begin{aligned} g_0^2&\equiv h_0^2+y_0^2\\ g_1&\equiv -2(h_0\phi_1+y_0\phi_2)w\\ g_2&\equiv w^2+l^2 \end{aligned}\right\} \tag{8}$$

g_0 is the distance from O_1 to the grating center.

Applying the reflection law on the surface of mirror M_1, we have for the direction cosines of reflected ray

$$\left.\begin{aligned} \lambda_2&=\lambda_1-2L(L\lambda_1+M\mu_1+N\nu_1)\\ \mu_2&=\mu_1-2M(L\lambda_1+M\mu_1+N\nu_1)\\ \nu_2&=\nu_1-2N(L\lambda_1+M\mu_1+N\nu_1) \end{aligned}\right\} \tag{9}$$

Using (2) (3) (4) and (5), we have from the first expression of (9)

$$x=\frac{R^2(\rho_1G_1+\rho_2s_1)}{R^2\rho_1-\rho_2(\rho_1^2-s^2)}$$

Let for convenience

$$\alpha\equiv\rho_2/\rho_1 \tag{10}$$

$$\varepsilon_1\equiv\rho_1^2-s^2 \tag{11}$$

$$a\equiv 1-(\alpha z_1/R^2) \tag{12}$$

Then,

$$x=(G_1+\alpha s_1)a^{-1} \tag{13}$$

Similarly, from the 2nd and 3rd expression of (9)

$$y=(G_2+\alpha s_2)a^{-1} \tag{14}$$

$$z=(l+\alpha z_1)a^{-1} \tag{15}$$

and hence from (3), (6) and (7)

$$\alpha^2 s^2 + 2\alpha k^2 + g^2 - R^2 a^2 = 0 \tag{16}$$

where,

$$k^2 \equiv s_1 G_1 + s_2 G_2 + z_1 l \tag{17}$$

From (1),

$$k^2 = k_0^2 + k_1 \tag{17'}$$

where,

$$k_0^2 \equiv h_0 s_1 - y_0 s_2 \tag{18}$$

$$k_1 \equiv (-s_1\phi_1 + s_2\phi_2)w + z_1 l \tag{19}$$

Solving (16) for α and remembering $\alpha > 0$, we obtain

$$\alpha = \frac{1}{s^2}\{-k^2 + Rsa(1-ma^{-2})^{\frac{1}{2}}\} \tag{20}$$

where,

$$m \equiv \frac{s^2 g^2 - k^4}{R^2 s^2} \tag{21}$$

m in the above expression is obtained from (7') and (17') or (8) and (19) as

$$m = m_0 + m_1 + m_2 \tag{22}$$

where,

$$\left.\begin{aligned} m_0 &\equiv \frac{s^2 g_0^2 - k_0^4}{R^2 s^2} \\ m_1 &\equiv \frac{s^2 g_1 - 2k_0^2 k_1}{R^2 s^2} = \frac{-2}{R^2 s^2}\{(\phi_1 m_{10} + \phi_2 m_{11})w + k_0^2 z_1 l\} \\ m_2 &\equiv \frac{s^2 g_2 - k_1^2}{R^2 s^2} \\ &= \frac{1}{R^2 s^2}\{(\phi_1^2 r_0^2 + \phi_2^2 m_{20} + 2\phi_1\phi_2 s_1 s_2)w^2 + 2(\phi_1 s_1 - \phi_2 s_2)w z_1 l + m_{21} l^2\} \end{aligned}\right\} \tag{23}$$

$$\left.\begin{aligned} m_{10} &\equiv h_0(s_2^2 + z_1^2) + s_1 s_2 y_0 \\ m_{11} &= h_0 s_1 s_2 + (s_1^2 + z_1^2) y_0 \\ r_0^2 &\equiv s_2^2 + z_1^2 \\ m_{20} &\equiv s_1^2 + z_1^2 \\ m_{21} &\equiv s_1^2 + s_2^2 \end{aligned}\right\} \tag{23'}$$

With the value of x, y and z, we obtain from (4), (5), (6), (7) and (17)

$$\rho_1^2 = R^2 + s^2 - 2(k^2 + \alpha s^2)a^{-1}$$

which becomes from (20),

$$\rho_1^2 = R^2 + s^2 - 2Rs(1-ma^{-2})^{\frac{1}{2}} \tag{24}$$

Similarly for the reflected ray, we have from (5)

$$\rho_2^2=R^2+g^2-2(g^2+\alpha k^2)a^{-1}$$

which becomes from (20)

$$\rho_2^2=R^2+g^2-2R^2ma^{-1}-2\frac{Rk^2}{s}(1-ma^{-2})^{\frac{1}{2}} \tag{25}$$

Since the factor a in the expressions (24) and (25) is an unknown quantity, it must be found. Using (12), (20), (11) and (24), we can obtain a relation for a as

$$a=1-\frac{1}{Rs^2}\{-k^2+Rsa(1-ma^{-2})^{\frac{1}{2}}\}\{R-2s(1-ma^{-2})^{\frac{1}{2}}\}$$

or

$$\left(1+\frac{k^2}{s^2}\right)-2ma^{-1}+a=\frac{1}{Rs}(R^2a+2k^2)(1-ma^{-2})^{\frac{1}{2}}$$

By squaring both side and by re-arranging

$$\varphi(a)\equiv\left(1-\frac{R^2}{s^2}\right)a^4+2\left(1-\frac{k^2}{s^2}\right)a^3+\left\{1-4\frac{g^2}{R^2}+\frac{1}{s^2}(2k^2+g^2)\right\}a^2$$

$$-4ma+4\frac{mg^2}{R^2}=0 \tag{26}$$

Since $a\approx 1$, approximate value of a is obtained from

$$a=1-\varphi(1)\Big/\frac{\partial\varphi(1)}{\partial a}$$

But since

$$\varphi(1)=4\left(1-\frac{g^2}{R^2}\right)\left(1-\frac{R^2}{4s^2}+m\right)$$

$$\frac{\partial\varphi(1)}{\partial a}=-6+2\left(1-\frac{g^2}{R^2}\right)\left(4-\frac{R^2}{s^2}\right)+2\left(1-\frac{k^2}{s^2}\right)-4m\approx-6$$

we obtain

$$a=1+\frac{2}{3}\left(1-\frac{g^2}{R^2}\right)\left(1-\frac{R^2}{4s^2}+m\right)+\cdots\cdots$$

By applying (7′) and (22) to the above expression, a can be expressed as

$$a=a_0+a_1+a_2+a_3+a_4+\cdots\cdots \tag{27}$$

where,

$$\left.\begin{aligned} a_0&\equiv 1+\frac{2}{3}p_0, \qquad p_0\equiv(C_0+m_0)D_0\\ a_1&\equiv\frac{2}{3}p_1, \qquad p_1\equiv D_0m_1-\frac{g_1}{R^2}(C_0+m_0)\end{aligned}\right\}$$

$$\left.\begin{aligned} &a_2\equiv\frac{2}{3}p_2\,, && p_2\equiv D_0m_2-\frac{1}{R^2}\{m_1g_1+g_2(C_0+m_0)\}\\ &a_3\equiv\frac{2}{3}p_3\,, && p_3\equiv-\frac{1}{R^2}(m_1g_2+m_2g_1)\\ &a_4\equiv\frac{2}{3}p_4\,, && p_4\equiv-\frac{1}{R^2}m_2g_2\\ &C_0\equiv1-\frac{R^2}{4s^2}\,, && D_0\equiv1-\frac{q_0^2}{R_2} \end{aligned}\right\} \tag{28}$$

These p_1, p_2 and p_3 are also expressed by (8) and (23) as follows.

$$\left.\begin{aligned} &p_1=\frac{2}{R^2}\{(\phi_1p_{10}+\phi_2p_{11})w-p_{12}z_1l\}\\ &p_{10}\equiv h_0(C_0-m_0)-\frac{m_{10}}{s^2}D_0\\ &p_{11}\equiv y_0(C_0-m_0)-\frac{m_{11}}{s^2}D_0\\ &p_{12}\equiv\frac{k_0^2}{s^2}D_0 \end{aligned}\right\} \tag{29}$$

$$\left.\begin{aligned} &p_2=\frac{1}{R^2}\{(\phi_1^2p_{20}+\phi_2^2p_{21}+2\phi_1\phi_2p_{22})w^2+(\phi_1p_{23}+\phi_2p_{24})wz_1l+p_{25}l^2\}\\ &p_{20}\equiv-(C_0+m_0)+4\frac{h_0m_{10}}{R^2s^2}+D_0\frac{r_0^2}{s^2}\\ &p_{21}\equiv-(C_0+m_0)+4\frac{y_0m_{11}}{R^2s^2}+D_0\frac{m_{20}}{s^2}\\ &p_{22}\equiv\frac{2}{R^2s^2}(h_0m_{11}+y_0m_{10})+D_0\frac{s_1s_2}{s^2}\\ &p_{23}\equiv4\frac{h_0k_0^2}{R^2s^2}+2D_0\frac{s_1}{s^2}\\ &p_{24}\equiv4\frac{y_0k_0^2}{R^2s^2}-2D_0\frac{s_2}{s^2}\\ &p_{25}\equiv-(C_0+m_0)+D_0\frac{m_{21}}{s^2} \end{aligned}\right\} \tag{30}$$

$$\left.\begin{aligned} &p_3=\frac{2}{R^4s^2}\{(\phi_1^3p_{30}+\phi_1^2\phi_2p_{31}+\phi_1\phi_2^2p_{32}+\phi_2^3p_{33})w^3\\ &\qquad+(\phi_1^2p_{34}+\phi_2^2p_{35}+\phi_1\phi_2p_{36})w^2z_1l+(\phi_1p_{37}+\phi_2p_{38})wl^2+p_{39}l^3\}\\ &p_{30}\equiv m_{10}+h_0r_0^2 \qquad p_{31}\equiv m_{11}+2h_0s_1s_2+y_0r_0^2\\ &p_{32}\equiv m_{10}+h_0m_{20}+2y_0s_1s_2 \qquad p_{33}\equiv m_{11}+y_0m_{20}\\ &p_{34}\equiv k_0^2+2h_0s_1 \qquad p_{35}\equiv k_0^2-2y_0s_2 \end{aligned}\right\} \tag{31}$$

$$p_{36}\equiv 2(s_1y_0-h_0s_2) \qquad p_{37}\equiv m_{10}+h_0m_{21}$$
$$p_{38}\equiv m_{11}+y_0m_{21} \qquad p_{39}\equiv z_1k_0^2$$

Therefore,

$$\left.\begin{aligned} a^{-2}&=a_0^{-2}(1+b_1+b_2+b_3+b_4+\cdots\cdots)\\ b_1&\equiv -2a_0^{-1}a_1\,, \qquad b_2\equiv -2a_0^{-1}a_2+3a_0^{-2}a_1^2\\ b_3&\equiv -2a_0^{-1}a_3+6a_0^{-2}a_1a_2-4a_0^{-3}a_1^3\\ b_4&\equiv -2a_0^{-1}a_4+3(a_0^{-2}a_2^2+2a_0^{-2}a_1a_3)-12a_0^{-3}a_1^2a_2+5a_0^{-4}a_1^4 \end{aligned}\right\} \tag{32}$$

$$\left.\begin{aligned} a^{-1}&=a_0^{-1}(1+b_1'+b_2'+b_3'+b_4'+\cdots\cdots)\\ b_1'&\equiv \frac{1}{2}\,b_1\,, \qquad b_2'\equiv \frac{1}{2}\,(b_2-a_0^{-2}a_1^2)\\ b_3'&\equiv \frac{1}{2}\,\{b_3-2a_0^{-3}a_1(a_2-a_1^2)\}\\ b_4'&\equiv \frac{1}{2}\,\{b_4-a_0^{-3}(a_2^2+2a_1a_3)+6a_0^{-3}a_1^2a_2-3a_0^{-4}a_1^4\} \end{aligned}\right\} \tag{33}$$

From (22) and (32), (24) becomes

$$\rho_1^2=R^2+s^2-2Rs\,[\beta_0-(m_1+m_0b_1)-(m_2+m_1b_1+m_0b_2)-(m_2b_1+m_1b_2+m_0b_3)$$
$$-(m_2b_2+m_1b_3+m_0b_4)-\cdots\cdots]^{\frac{1}{2}} \tag{33'}$$

where

$$\beta_0\equiv 1-m_0a_0^{-2} \tag{34}$$

Then ρ_1^2 can be rewritten by expanding the last term of (33′) in series as

$$\rho_1^2=\rho_{10}^2+\rho_{11}'+\rho_{12}'+\rho_{13}'+\rho_{14}'+\cdots\cdots \tag{35}$$

$$\rho_{10}^2\equiv R^2+s^2-2Rs\beta_0^{\frac{1}{2}} \tag{36}$$

$$\rho_{11}'\equiv Rsa_0^{-2}\beta_0^{-\frac{1}{2}}(m_1+m_0b_1) \tag{37}$$

$$\rho_{12}'\equiv Rsa_0^{-2}\beta_0^{-\frac{1}{2}}\left\{m_2+m_1b_1+m_0b_2+\frac{1}{4}\,a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)^2\right\} \tag{38}$$

$$\rho_{13}'\equiv Rsa_0^{-2}\beta_0^{-\frac{1}{2}}\left\{m_2b_1+m_1b_2+m_0b_3+\frac{1}{2}\,a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)(m_2+m_1b_1+m_0b_2)\right.$$
$$\left.+\frac{1}{8}\,a_0^{-4}\beta_0^{-2}(m_1+m_0b_1)^3\right\} \tag{39}$$

$$\rho_{14}'\equiv Rsa_0^{-2}\beta_0^{-\frac{1}{2}}(m_2b_2+m_1b_3+m_0b_4)+\frac{\beta_0^{-\frac{1}{2}}}{4Rs}(\rho_{12}'^2+2\rho_{11}'\rho_{13}') \tag{40}$$

Using (23), (23′), (29) and 30), we have ρ_{11}' and ρ_{12}' explicitly in terms of w and l, namely

$$\left.\begin{aligned}
&\rho_{11}' = \frac{-2}{Rs}\{(\phi_1 A_{10}+\phi_2 A_{11})w + A_{12} z_1 l\} \\
&A_{10} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(m_{10} + \frac{4}{3}s^2 m_0 p_{10} a_0^{-1}\right) \\
&A_{11} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(m_{11} + \frac{4}{3}s^2 m_0 p_{11} a_0^{-1}\right) \\
&A_{12} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(k_0^2 - \frac{4}{3}s^2 m_0 p_{18} a_0^{-1}\right)
\end{aligned}\right\} \quad (37')$$

$$\left.\begin{aligned}
&\rho_{12}' = (\phi_1^2 A_{20}+\phi_2^2 A_{21}+2\phi_1\phi_2 A_{22})w^2 + (\phi_1 A_{23}+\phi_2 A_{24})2wz_1 l + A_{25} l^2 \\
&A_{20} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(\frac{r_0^2}{Rs} + \frac{16a_0^{-1}}{3R^3 s} m_{10} p_{10} - \frac{4sa_0^{-1}}{3R} m_0 p_{20} \right. \\
&\qquad\qquad \left. + \frac{16sa_6^{-2}}{3R^3} m_0 p_{10}^2 + \frac{\beta_0^{-\frac{1}{2}}}{R^3 s^3} A_{10}^2\right) \\
&A_{21} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(\frac{m_{20}}{Rs} + \frac{16a_0^{-1}}{3R^3 s} m_{11} p_{11} - \frac{4sa_0^{-1}}{3R} m_0 p_{21} \right. \\
&\qquad\qquad \left. + \frac{16sa_0^{-2}}{3R^3} m_0 p_{11}^2 + \frac{\beta_0^{-\frac{1}{2}}}{R^3 s^3} A_{11}^2\right) \\
&A_{22} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left\{\frac{s_1 s_2}{Rs} + \frac{8a_0^{-1}}{3R^3 s}(p_{10}m_{11}+p_{11}m_{10}) + \cdots\cdots\right\} \\
&A_{23} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left\{\frac{s_1}{Rs} + \frac{8a_0^{-1}}{3R^3 s}(k_0^2 p_{10} - m_{10}p_{12}) + \cdots\cdots\right\} \\
&A_{24} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left\{-\frac{s_2}{Rs} + \frac{8a_0^{-1}}{3R^3 s}(k_0^2 p_{11} - m_{11}p_{10}) + \cdots\cdots\right\} \\
&A_{25} \equiv a_0^{-2}\beta_0^{-\frac{1}{2}}\left(\frac{m_{21}}{Rs} - \frac{16a_6^{-2}}{3R^3 s} k_0^2 p_{12} z_1^2 - \frac{4sa_0^{-1}}{3R} m_0 p_{25} \right. \\
&\qquad\qquad \left. + \frac{16sa_0^{-2}}{3R^3} m_0 p_{12}^2 z_1^2 + \frac{\beta_0^{-\frac{1}{2}}}{R^3 s^3} A_{12}^2 z_1^2\right)
\end{aligned}\right\} \quad (38')$$

By further series expansion, incident light path ρ_1 is finally obtained as

$$\rho_1 = \rho_{10}+\rho_{11}+\rho_{12}+\rho_{13}+\rho_{14}+\cdots\cdots \quad (41)$$

$$\rho_{11} \equiv \frac{1}{2\rho_{10}}\rho_{11}' \quad (42)$$

$$\rho_{12} \equiv \frac{1}{2\rho_{10}}\left(\rho_{12}' - \frac{1}{4\rho_{10}^2}\rho_{11}'^2\right) \quad (43)$$

$$\rho_{13} \equiv \frac{1}{2\rho_{10}}\left(\rho_{13}' - \frac{1}{2\rho_{10}^2}\rho_{11}'\rho_{12}' + \frac{1}{8\rho_{10}^4}\rho_{11}'^3\right) \quad (44)$$

$$\rho_{14}\equiv\frac{1}{2\rho_{10}}\left\{\rho_{14}'-\frac{1}{4\rho_{10}^{2}}(2\rho_{11}'\rho_{13}'+\rho_{12}'^{2})+\frac{3}{8\rho_{10}^{4}}\rho_{11}'^{2}\rho_{12}'-\frac{5}{64\rho_{10}^{6}}\rho_{11}'^{4}\right\} \quad (45)$$

ρ_{10} is the path length for the incident ray to reach the grating center.

In an exactly similar way, reflected light path ρ_2 becomes

$$\rho_2^{2}=R^{2}+g^{2}-2R^{2}a_0^{-1}\{m_0+(m_1+m_0b_1')+(m_2+m_1b_1'+m_0b_2')+(m_2b_1'+m_1b_2'+m_0b_3')+(m_2b_2'+m_1b_3'+m_0b_4')+\cdots\cdots\}$$

$$-2\frac{Rk_0^{2}}{s}\{\beta_0-a_0^{-2}(m_1+m_0b_1)-a_0^{-2}(m_2+m_1b_1+m_0b_2)-a_0^{-2}(m_2b_1+m_1b_2+m_0b_3)-a_0^{-2}(m_2b_2+m_1b_3+m_0b_4)-\cdots\cdots\}^{\frac{1}{2}}$$

By series expansion,

$$\rho_2^{2}=\rho_{20}^{2}+\rho_{21}'+\rho_{22}'+\rho_{23}'+\rho_{24}'+\cdots\cdots \quad (46)$$

$$\rho_{20}^{2}\equiv R^{2}+g_0^{2}-2R^{2}m_0a_0^{-1}-2\frac{Rk_0^{2}}{s}\beta_0^{\frac{1}{2}} \quad (47)$$

$$\rho_{21}'\equiv g_1-2R^{2}a_0^{-1}(m_1+m_0b_1'+\frac{Rk_0^{2}}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}(m_1+m_0b_1)-2\frac{Rk_1}{s}\beta_0^{\frac{1}{2}} \quad (48)$$

$$\rho_{22}'\equiv g_2-2R^{2}a_0^{-1}(m_2+m_1b_1'+m_0b_2')+\frac{Rk_0^{2}}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}\left\{m_2+m_1b_1+m_0b_2+\frac{1}{4}a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)^{2}\right\}+\frac{R}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}k_1(m_1+m_0b_1) \quad (49)$$

$$\rho_{23}'\equiv-2R^{2}a_0^{-1}(m_3b_1'+m_1b_2'+m_0b_3')+\frac{Rk_0^{2}}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}\left\{m_2b_1+m_1b_2+m_0b_3+\frac{1}{2}a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)(m_2+m_1b_1+m_0b_2)+\frac{1}{8}a_0^{-4}\beta_0^{-2}(m_1+m_0b_1)^{3}\right\}+\frac{R}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}k_1\left\{m_2+m_1b_1+m_0b_2+\frac{1}{4}a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)^{2}\right\} \quad (50)$$

$$\rho_{24}'\equiv-2R^{2}a_0^{-1}(m_2b_2'+m_1b_3'+m_0b_4')+\frac{Rk_0^{2}}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}\left[m_2b_2+m_1b_3+m_0b_4+\frac{1}{4}a_0^{-2}\beta_0^{-1}\{2(m_1+m_0b_1)(m_2b_1+m_1b_2+m_0b_3)+(m_2+m_1b_1+m_0b_2)^{2}\}+\frac{3}{8}a_0^{-4}\beta_0^{-2}(m_1+m_0b_1)^{2}(m_2+m_1b_1+m_0b_2)+\frac{5}{64}a_0^{-6}\beta_0^{-3}(m_1+m_0b_1)^{4}\right]+\frac{R}{s}a_0^{-2}\beta_0^{-\frac{1}{2}}k_1\left\{m_2b_1+m_1b_2+m_0b_3+\frac{1}{2}a_0^{-2}\beta_0^{-1}(m_1+m_0b_1)(m_2+m_1b_1+m_0b_2)+\frac{1}{8}a_0^{-4}\beta_0^{-2}(m_1+m_0b_1)^{3}\right\} \quad (51)$$

From (32), (33) and further from (23) and (37), we obtain for ρ_{21}', ρ_{22}' and ρ_{23}'

$$\rho_{21}'=g_1(1-a_0^{-1})-\frac{1}{s}\beta_0^{\frac{1}{2}}(2k_1+a_0\rho_{11}')\left(R-\frac{k_0^2}{s}a_0^{-1}\beta_0^{-\frac{1}{2}}\right)$$

By putting

$$Q_0\equiv\frac{1}{\rho_{10}}-\frac{a_0\beta_0^{\frac{1}{2}}}{\rho_{20}s}\left(R-\frac{k_0^2}{s}a_0^{-1}\beta_0^{-\frac{1}{2}}\right) \tag{52}$$

or

$$R-\frac{k_0^2}{s}a_0^{-1}\beta_0^{-\frac{1}{2}}=\left(\frac{1}{\rho_{10}}-Q_0\right)\rho_{20}sa_0^{-1}\beta_0^{-\frac{1}{2}} \tag{52'}$$

ρ_{21}' becomes

$$\rho_{21}'=-\frac{\rho_{20}}{\rho_{10}}\rho_{11}'+B_1+B_1'Q_0 \tag{53}$$

$$\left.\begin{aligned}B_1&\equiv g_1(1-a_0^{-1})-2\frac{\rho_{20}}{\rho_{10}}a_0^{-1}k_1\\ B_1'&\equiv\rho_{20}a_0^{-1}(2k_1+a_0\rho_{11}')\end{aligned}\right\} \tag{54}$$

From (8), (19) and (37′)

$$\left.\begin{aligned}B_1+B_1'Q_0=-2[\{\phi_1(B_{10}+B_{10}'Q_0)+\phi_3(B_{11}+B_{11}'Q_0)\}w+(B_{12}+B_{12}'Q_0)z_1l]\\ B_{10}\equiv h_0(1-a_0^{-1})-\frac{\rho_{20}}{\rho_{10}}s_1a_0^{-1}\qquad B_{10}'\equiv\rho_{20}a_0^{-1}\left(s_1+\frac{a_0}{Rs}A_{10}\right)\\ B_{11}\equiv\gamma_0(1-a_0^{-1})+\frac{\rho_{20}}{\rho_{10}}s_2a_0^{-1}\qquad B_{11}'\equiv\rho_{20}a_0^{-1}\left(-s_2+\frac{a_0}{Rs}A_{11}\right)\\ B_{12}\equiv\frac{\rho_{20}}{\rho_{10}}a_0^{-1}\qquad B_{12}'\equiv\rho_{20}a_0^{-1}\left(-1+\frac{a_0}{Rs}A_{12}\right)\end{aligned}\right\} \tag{54'}$$

Using (33), (37) and (38), we have for ρ_{22}'

$$\rho_{22}'=g_2-R^2m_2a_0^{-1}-\frac{\rho_{12}'}{s}a_0\beta_0^{\frac{1}{2}}\left(R-\frac{k_0^2}{s}a_0^{-1}\beta_0^{-\frac{1}{2}}\right)$$
$$+\frac{a_0}{4s^2}\rho_{11}'^2+\frac{k_1}{s^2}\rho_{11}'+R^2m_0a_0^{-3}a_1^2$$

Further from the relations (23) and (52′)

$$\rho_{22}'=-\frac{\rho_{20}}{\rho_{10}}\rho_{12}'+B_2+B_2'Q_0 \tag{55}$$

$$\left.\begin{aligned}B_2&\equiv g_2(1-a_0^{-1})+\frac{a_0^{-1}}{s^2}\left(k_1+\frac{a_0}{2}\rho_{11}'\right)^2+\frac{4}{9}R^2a_0^{-3}m_0p_1^2\\ B_2'&\equiv\rho_{12}'\rho_{20}a_0^{-1}\end{aligned}\right\} \tag{56}$$

From (8), (19), (29), (37′) and (33′),

$$
\left.
\begin{aligned}
B_2+B_2'Q_0 &= \{\phi_1^2(B_{20}+B_{20}'Q_0)+\phi_2^2(B_{21}+B_{21}'Q_0)+\phi_1\phi_2(2B_{22}+B_{22}'Q_0)\}\,w^2 \\
&\quad +\{\phi_1(-2B_{23}+B_{23}'Q_0)+\phi_2(2B_{24}+B_{24}'Q_0)\}\,wz_1l+(B_{25}+B_{25}'Q_0)l^2 \\
B_{20} &\equiv 1-a_0^{-1}+\frac{a_0^{-1}}{s^2}\left(s_1+\frac{a_0}{Rs}A_{10}\right)^2+\frac{16a_0^{-3}}{9R^2}m_0p_{10}^2 \\
B_{21} &\equiv 1-a_0^{-1}+\frac{a_0^{-1}}{s^2}\left(s_2-\frac{a_0}{Rs}A_{11}\right)^2+\frac{16a_0^{-3}}{9R^2}m_0p_{11}^2 \\
B_{22} &\equiv -\frac{a_0^{-1}}{s^2}\left(s_1+\frac{a_0}{Rs}A_{10}\right)\left(s_2-\frac{a_0}{Rs}A_{11}\right)+\frac{16a_0^{-3}}{9R^2}m_0p_{10}p_{11} \\
B_{23} &\equiv \frac{a_0^{-1}}{s^2}\left(1-\frac{a_0}{Rs}A_{12}\right)\left(s_1+\frac{a_0}{Rs}A_{10}\right)+\cdots\cdots \\
B_{24} &\equiv \frac{a_0^{-1}}{s^2}\left(1-\frac{a_0}{Rs}A_{12}\right)\left(s_2-\frac{a_0}{Rs}A_{11}\right)-\frac{16a_0^{-3}}{9R^2}m_0p_{11}p_{12} \\
B_{25} &\equiv 1-a_0^{-1}+\frac{a_0^{-1}}{s^2}\left(1-\frac{a_0}{Rs}A_{12}\right)^2z_1^2+\frac{16a_0^{-3}}{9R^2}m_0p_{12}^2z_1^2 \\
B_{20}' &\equiv \rho_{20}a_0^{-1}A_{20}\,, \qquad B_{21}' \equiv \rho_{20}a_0^{-1}A_{21}\,, \qquad B_{22}' \equiv \rho_{20}a_0^{-1}A_{22} \\
B_{23}' &\equiv \rho_{20}a_0^{-1}A_{23}\,, \qquad B_{24}' \equiv \rho_{20}a_0^{-1}A_{24}\,, \qquad B_{25}' \equiv \rho_{20}a_0^{-1}A_{25}
\end{aligned}
\right\} \quad (56)'
$$

The second term of ρ_{23}' (50) becomes $(k^2/s^2)\,\rho_{13}'$ by the use of (39). Therefore from (28), (32) and (33), ρ_{23}' is finally expressed as follows.

$$\rho_{23}' = -\frac{\rho_{20}}{\rho_{10}}\rho_{13}' + B_3 + \rho_{20}\rho_{13}'Q_0 \tag{57}$$

$$
\begin{aligned}
B_3 \equiv{}& \frac{a_0}{2s^2}\rho_{11}'\rho_{12}' + \frac{k_1}{s^2}\rho_{12}' + \frac{4}{9}R^2a_0^{-3}m_1p_1^2 \\
&+ \frac{8}{9}R^2m_0a_0^{-3}p_1p_2 - \frac{16}{27}R^2m_0a_0^{-4}p_1^3
\end{aligned} \tag{58}
$$

or from (19), (23′), (29), (30), (37′) and (38′)

$$
\left.
\begin{aligned}
B_3 &= (\phi_1^3B_{30}+\phi_1^2\phi_2B_{31}+\phi_1\phi_2^2B_{32}+\phi_2^3B_{33})w^3 \\
&\quad +(\phi_1^2B_{34}+\phi_2^2B_{35}+\phi_1\phi_2B_{36})w^2z_1l+(\phi_1B_{37}+\phi_2B_{38})wl^2+B_{39}z_1l^3 \\
B_{30} &\equiv -\frac{a_0}{Rs^3}A_{10}A_{20}-\frac{1}{s^2}s_1A_{20}-\frac{32}{9R^4s^2}a_0^{-3}m_{10}p_{10}^2 \\
&\quad +\frac{16}{9R^2}m_0a_0^{-3}p_{10}p_{20}-\frac{128}{27}m_0a_0^{-4}p_{10}^3 \\
B_{31} &\equiv -\frac{a_0}{Rs^3}(A_{10}A_{22}+A_{11}A_{20})-\frac{1}{s^2}(s_1A_{22}-s_2A_{20})+\cdots\cdots \\
B_{32} &\equiv -\frac{a_0}{Rs^3}(A_{10}A_{21}+A_{11}A_{22})-\frac{1}{s^2}(s_1A_{21}-s_2A_{22})
\end{aligned}
\right\}
$$

$$
\left.
\begin{aligned}
&-\frac{32a_0^{-3}}{9R^4s^3}(m_{10}p_{11}^2+2m_{11}p_{10}p_{11})+\frac{16}{9R^2}m_0a_0^{-3}(p_{10}p_{21}+p_{11}p_{22})\\
&-\frac{128}{9}m_0a_0^{-4}p_{10}p_{11}^2\\
B_{33}\equiv&-\frac{a_0}{Rs^3}A_{11}A_{21}+\frac{1}{s^2}s_2A_{21}+\cdots\cdots\\
B_{34}\equiv&-\frac{a_0}{Rs^3}(A_{12}A_{20}+A_{10}A_{23})+\cdots\cdots\\
B_{35}\equiv&-\frac{a_0}{Rs^3}(A_{12}A_{21}+A_{11}A_{24})+\cdots\cdots\\
&\cdots\\
B_{39}\equiv&-\frac{a_0}{Rs^3}A_{12}A_{25}+\frac{1}{s^2}A_{25}-\frac{32a_0^{-3}}{9R^4s^2}k_0^2p_{12}^2z_1^2\\
&-\frac{16a_0^{-3}}{9R_2}m_0p_{12}p_{25}+4\frac{a_0^{-4}}{R^4}m_0p_{12}^2z_1^2
\end{aligned}
\right\} \quad (58')
$$

Similarly,

$$\rho_{24}'=-\frac{\rho_{20}}{\rho_{10}}\rho_{14}'+B_4+\rho_{20}\rho_{14}'Q_0 \quad (59)$$

$$
\begin{aligned}
B_4\equiv&\frac{a_0}{4s^2}(\rho_{12}'^2+2\rho_{11}'\rho_{13}')+\frac{k_1}{s^2}\rho_{13}'+\frac{4}{9}R^2a_0^{-3}m_2p_1^2+\frac{8}{9}R^2a_0^{-3}m_1p_1p_2\\
&-\frac{16}{27}R^2a_0^{-4}m_1p_1^3+\frac{4}{9}R^2m_0a_0^{-4}(p_2^2+2p_1p_3)-\frac{16}{9}R^2m_0a_0^{-4}p_1^2p_2\\
&+\frac{16}{27}R^2m_0a_0^{-5}p_1^4
\end{aligned} \quad (60)
$$

By further series expansion, ρ_2 is finally presented as follows

$$\rho_2=\rho_{20}+\rho_{21}+\rho_{22}+\rho_{23}+\rho_{24}+\cdots\cdots \quad (61)$$

$$\rho_{21}\equiv\frac{1}{2\rho_{20}}\rho_{21}' \quad (62)$$

$$\rho_{22}\equiv\frac{1}{2\rho_{20}}\left(\rho_{22}'-\frac{1}{4\rho_{20}^2}\rho_{21}'^2\right) \quad (63)$$

$$\rho_{23}\equiv\frac{1}{2\rho_{20}}\left(\rho_{23}'-\frac{1}{2\rho_{20}^2}\rho_{21}'\rho_{22}'+\frac{1}{8\rho_{20}^4}\rho_{21}'^3\right) \quad (64)$$

$$\rho_{24}\equiv\frac{1}{2\rho_{20}}\left\{\rho_{24}'-\frac{1}{4\rho_{20}^2}(2\rho_{31}'\rho_{23}'+\rho_{22}'^2)+\frac{3}{8\rho_{20}^4}\rho_{21}'^2\rho_{22}'-\frac{5}{64\rho_{20}^6}\rho_{21}'^4\right\} \quad (65)$$

Therefore, optical path traversed through M_1-system is

$$\rho_1+\rho_2=[\rho_{10}+\rho_{20}]+\frac{1}{2}\left[\frac{1}{\rho_{10}}\rho_{11}'+\frac{1}{\rho_{20}}\rho_{21}'\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{10}}\rho_{12}'+\frac{1}{\rho_{20}}\rho_{23}'-\frac{1}{4\rho_{10}^3}\rho_{11}'^2-\frac{1}{4\rho_{20}^3}\rho_{21}'^2\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{10}}\rho_{13}'+\frac{1}{\rho_{20}}\rho_{23}'-\frac{1}{2\rho_{20}^2}\rho_{21}'\left(\frac{1}{\rho_{10}}\rho_{12}'+\frac{1}{\rho_{20}}\rho_{22}'-\frac{1}{4\rho_{10}^3}\rho_{11}'^2-\frac{1}{4\rho_{20}^3}\rho_{21}'^2\right)\right.$$
$$\left.-\frac{1}{2\rho_{10}}\rho_{12}'\left(\frac{1}{\rho_{10}^2}\rho_{11}'-\frac{1}{\rho_{20}^2}\rho_{21}'\right)+\frac{1}{8\rho_{10}^3}\rho_{11}'^2\left(\frac{1}{\rho_{10}^2}\rho_{11}'-\frac{1}{\rho_{20}^2}\rho_{21}'\right)\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{10}}\rho_{14}'+\frac{1}{\rho_{20}}\rho_{24}'-\frac{1}{2\rho_{20}^2}\rho_{21}'\left\{\frac{1}{\rho_{10}}\rho_{13}'+\frac{1}{\rho_2^0}\rho_{23}'-\frac{1}{2\rho_{20}^2}\rho_{21}'\right.\right.$$
$$\left(\frac{1}{\rho_{10}}\rho_{12}'+\frac{1}{\rho_{20}}\rho_{22}'-\frac{1}{4\rho_{10}^3}\rho_{11}'^2-\frac{1}{4\rho_{20}^3}\rho_{21}'^2\right)-\frac{1}{2\rho_{10}}\left(\rho_{12}'-\frac{1}{4\rho_{10}^2}\rho_{11}'^2\right)$$
$$\left.\left(\frac{1}{\rho_{10}^2}\rho_{11}'-\frac{1}{\rho_{20}^2}\rho_{21}'\right)\right\}-\frac{1}{2\rho_{10}}\left(\rho_{13}'+\frac{1}{2\rho_{10}^2}\rho_{11}'\rho_{12}'\right)\left(\frac{1}{\rho_{10}^2}\rho_{11}'-\frac{1}{\rho_{20}^2}\rho_{21}'\right)$$
$$\left.+\frac{1}{8}\left(\frac{1}{\rho_{10}^5}\rho_{11}'^2\rho_{12}'+\frac{1}{\rho_{20}^5}\rho_{21}'^2\rho_{22}'-\frac{1}{8\rho_{20}^7}\rho_{21}'^4\right)\right]+\cdots\cdots$$

Inserting (53), (55), (57) and (59) into the above expression,

$$\rho_1+\rho_2=[\rho_{10}+\rho_{20}]+\frac{1}{2}\left[\frac{1}{\rho_{20}}(B_1+B_1'Q_0)\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{20}}(B_2+B_2'Q_0)-\frac{1}{4\rho_{10}^3}\rho_{11}'^2-\frac{1}{4\rho_{20}}\left\{\frac{1}{\rho_{10}}\rho_{11}'-\frac{1}{\rho_{20}}(B_1+B_1'Q_0\right\}^2\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{20}}B_3+\frac{1}{2\rho_{20}}\left(\frac{1}{\rho_{10}}\rho_{11}'-\frac{1}{\rho_{20}}B_1\right)\left\{\frac{1}{\rho_{20}}B_2-\frac{1}{4\rho_{10}^3}\rho_{11}'^2-\frac{1}{4\rho_{20}}\right.\right.$$
$$\left.\left(\frac{1}{\rho_{10}}\rho_{11}'-\frac{1}{\rho_{20}}B_1\right)^2\right\}-\frac{1}{2\rho_{10}}\left(\rho_{12}'-\frac{1}{4\rho_{10}^2}\rho_{11}'^2\right)$$
$$\left.\left\{\frac{1}{\rho_{10}}\left(\frac{1}{\rho_{10}}+\frac{1}{\rho_{20}}\right)\rho_{11}'-\frac{1}{\rho_{20}^2}B_1\right\}\right]$$
$$+\frac{1}{2}\left[\frac{1}{\rho_{20}}B_4+\frac{1}{2\rho_{20}}\left(\frac{1}{\rho_{10}}\rho_{11}'-\frac{1}{\rho_{20}}B_1\right)\left\{\frac{1}{\rho_{20}}B_3+\cdots\cdots\right\}\right.$$
$$-\frac{1}{2\rho_{10}}\left(\rho_{13}'+\frac{1}{2\rho_{10}^2}\rho_{11}'\rho_{12}'\right)\left(\frac{1}{\rho_{10}^2}\rho_{11}'+\cdots\cdots\right)$$
$$+\frac{1}{8}\left\{\frac{1}{\rho_{10}^5}\rho_{11}'^2\rho_{12}'-\frac{1}{\rho_{20}^2}\left(\frac{1}{\rho_{10}}\rho_{12}'+B_2\right)\left(\frac{1}{\rho_{10}^2}\rho_{11}'^2+\cdots\cdots\right)\right.$$
$$\left.\left.+\frac{1}{8\rho_{20}^3}\left(\frac{1}{\rho_{20}^2}\rho_{11}'^2+\cdots\cdots\right)^3\right\}\right]$$
$$+\cdots\cdots \qquad (66)$$

ρ_{20} is the path length of reflected ray to reach the grating center. As will be mentioned later, $Q_0=0$ (52) is approximately the condition for image formation, therefore Q_0 is made zero in third and fourth terms. Furthermore, from (37′), (38′) (54′), (56′) and (58′), $(\rho_1+\rho_2)$ is explicitly expressed in terms of w and l as follows.

$$\rho_1+\rho_2=[\rho_{10}+\rho_{20}]+\frac{1}{2}[(\phi_1\psi_{10}+\phi_2\psi_{11})w+\psi_{12}z_1l]$$
$$+\frac{1}{2}[(\phi_1^2\psi_{20}+\phi_2^2\psi_{21}+\phi_1\phi_2\psi_{22})w^2+(\phi_1\psi_{23}+\phi_2\psi_{24})wz_1l+\psi_{25}l^2]$$
$$+\frac{1}{2}[(\phi_1^3\psi_{30}+\phi_1^2\phi_2\psi_{31}+\phi_1\phi_2^2\psi_{32}+\phi_2^3\psi_{33})w^3$$
$$+(\phi_1^2\psi_{34}+\phi_2^2\psi_{35}+\phi_1\phi_2\psi_{36})w^2z_1l+(\phi_1\psi_{37}+\phi_2\psi_{38})wl^2+\psi_{39}z_1l^3]$$
$$+\frac{1}{2}[\psi_{40}(w^2+l^2)^2+\cdots\cdots]+\cdots\cdots \qquad (66')$$

For M_2-system, an analogous expression for $(\rho_2'+\rho_1')$ may be derived by the following transformation of coordinates.

$$s_1\to s_1'\equiv x_1'-x_0', \qquad s_2\to s_2'\equiv y_1'-y_0'$$
$$z_1\to z_1', \qquad R\to R'$$

and

$$\rho_{11}'\to\rho_{11}'', \quad \rho_{12}'\to\rho_{12}'', \quad \rho_{13}'\to\rho_{13}'', \quad \rho_{14}'\to\rho_{14}''$$
$$B_1\to E_1, \quad B_1'\to E_1', \quad B_2\to E_2, \quad \cdots\cdots$$

Therefore, the total light path length or optical path function L is

$$L=\rho_1+\rho_2+\rho_2'+\rho_1'+\frac{n\lambda}{d}w \qquad (67)$$

where n is the order number of spectrum, d the grating constant.

If L is constant independently of w and l, the Fermat's principle will be satisfied. That is,

$$\frac{\partial L}{\partial w}=0, \qquad \frac{\partial L}{\partial l}=0 \qquad (68)$$

This means that the optical system is free from aberration. However, it is generally difficult to satisfy the above two conditions at the same time, therefore, the system is not wholly devoid of aberration. The residual aberration, however, can be made considerably small by proper alignment of optical parts. To deal with this aberrations, it is convenient to devide it into various known aberrations.

Therefore, following Beutler's method, let

$$L=\left(F_1+F_1'+\frac{n\lambda}{d}w\right)+(F_2+F_2')+(F_3+F_3')+(F_4+F_4')+\cdots\cdots \qquad (69)$$

$$F_1\equiv\rho_{10}+\rho_{20}+\frac{1}{2}(\phi_1\psi_{10}+\phi_2\psi_{11})w \qquad (70)$$

$$F_2 \equiv \frac{1}{2}(\phi_1^2\psi_{20}+\phi_2^2\psi_{21}+\phi_1\phi_2\psi_{22})w^2 + \frac{1}{2}(\phi_1^3\psi_{30}+\phi_1^2\phi_2\psi_{31}+\phi_1\phi_2^2\psi_{32}+\phi_2^3\psi_{33})w^3 \tag{71}$$

$$F_3 \equiv \frac{1}{2}\psi_{12}z_1 l + \frac{1}{2}\psi_{20}l^2 + \frac{1}{2}\psi_{39}z_1 l^3 \tag{72}$$

$$F_4 \equiv \frac{1}{2}\psi_{40}(w^2+l^2)^2 \tag{73}$$

where, ψ_{10}, ψ_{11}, ... are expressed as follows by referring to (66), (66′) and from (37′), (38′), (54′), (56′) and (58′)

$$\psi_{10} \equiv -\frac{2}{\rho_{20}}(B_{10}+B_{10}'Q_0) \tag{74}$$

$$\psi_{11} \equiv -\frac{2}{\rho_{20}}(B_{11}+B_{11}'Q_0) \tag{75}$$

$$\psi_{12} \equiv -\frac{2}{\rho_{20}}(B_{12}+B_{12}'Q_0) \tag{76}$$

$$\psi_{20} \equiv \frac{1}{\rho_{20}}(B_{20}+B_{20}'Q_0) - \frac{1}{\rho_{10}^3R^2s^2}A_{10}^2 - \frac{1}{\rho_{20}}\left\{\frac{1}{\rho_{20}}(B_{10}+B_{10}'Q_0)+\frac{1}{\rho_{10}Rs}A_{10}\right\}^2 \tag{77}$$

$$\psi_{21} \equiv \frac{1}{\rho_{20}}(B_{21}+B_{21}'Q_0) - \frac{1}{\rho_{10}^3R^2s^2}A_{11}^2 - \frac{1}{\rho_{20}}\left\{\frac{1}{\rho_{20}}(B_{11}+B_{11}'Q_0)+\frac{1}{\rho_{10}Rs}A_{11}\right\}^2 \tag{78}$$

$$\psi_{22} \equiv \frac{1}{\rho_{20}}(B_{22}+B_{22}'Q_0) - \frac{2}{\rho_{10}^3R^2s^2}A_{10}A_{11} + \cdots\cdots \tag{79}$$

$$\psi_{25} \equiv \frac{1}{\rho_{20}}(B_{25}+B_{25}'Q_0) - \frac{1}{\rho_{10}^3R^2s^2}A_{12}^2z_1^2 - \frac{1}{\rho_{20}}\left\{\frac{1}{\rho_{20}}(B_{12}+B_{12}'Q_0)+\frac{1}{\rho_{10}Rs}A_{12}\right\}^2 z_1^2 \tag{80}$$

$$\psi_{30} \equiv \frac{1}{\rho_{20}}B_{30} - \frac{1}{\rho_{20}}\psi_{20}\left(\frac{A_{10}}{\rho_{10}Rs}, \frac{B_{10}}{\rho_{20}}\right) + \frac{1}{\rho_{10}}\left(A_{20}-\frac{A_{10}^2}{\rho_{10}^2R^2s^2}\right)\left\{\frac{A_{10}}{\rho_{10}Rs}\left(\frac{1}{\rho_{10}}+\frac{1}{\rho_{20}}\right)-\frac{B_{10}}{\rho_{20}^2}\right\} \tag{81}$$

$$\psi_{31} \equiv \frac{1}{\rho_{20}}B_{31} + \cdots\cdots \tag{82}$$

$$\psi_{32} \equiv \frac{1}{\rho_{20}}B_{32} - \frac{1}{\rho_{20}}\left\{\psi_{21}\left(\frac{A_{10}}{\rho_{10}Rs}-\frac{B_{10}}{\rho_{20}}\right)+\psi_{22}\left(\frac{A_{11}}{\rho_{10}Rs}-\frac{B_{11}}{\rho_{20}}\right)\right\} + \frac{1}{\rho_{10}}\left[\left(A_{21}-\frac{A_{11}^2}{\rho_{10}^2R^2s^2}\right)\left\{\frac{A_{10}}{\rho_{10}Rs}\left(\frac{1}{\rho_{10}}+\frac{1}{\rho_{20}}\right)-\frac{B_{10}}{\rho_{20}^2}\right\} + \left(A_{20}-\frac{A_{10}^2}{\rho_{10}^2R^2s^2}\right)\left\{\frac{A_{11}}{\rho_{10}Rs}\left(\frac{1}{\rho_{10}}+\frac{1}{\rho_{20}}\right)-\frac{B_{11}}{\rho_{20}^2}\right\}\right] \tag{83}$$

$$\psi_{33}\equiv\frac{1}{\rho_{20}}B_{33}-\frac{\psi_{21}}{\rho_{20}}\left(\frac{A_{11}}{\rho_{10}Rs}-\frac{B_{11}}{\rho_{20}}\right)+\cdots\cdots \tag{84}$$

$$\psi_{39}\equiv\frac{1}{\rho_{20}}B_{39}-\frac{\psi_{25}}{\rho_{20}}\left(\frac{A_{12}}{\rho_{10}Rs}-\frac{B_{12}}{\rho_{20}}\right)+\cdots\cdots \tag{85}$$

Lastly, it is clear from (66) that the term which includes $g_2^2=(w^2+l^2)^2$ in ψ_{40} is only $(1/\rho_{20})\ B_4$, therefore the term which includes only g_2^2 in (60) is

$$\frac{a_0}{4s^2}\rho_{12}'^2+\frac{4}{9}R^2m_0a_0^{-4}p_2^2$$

Further, from (38) and (28) the above becomes

$$\frac{a_0}{4s^2}\{Rsa_0^{-2}\beta_0^{-\frac{1}{2}}(m_2+m_0b_2)\}^2+\frac{4}{9}R^2m_0a_0^{-4}\left\{D_0m_2-\frac{g_2}{R^2}(C_0+m_0)\right\}^2$$

Since the terms which have g_2 in m_2 (23) and b_2 (32) are g_2/R^2 and $2a_0^{-1}a_2\left(=\frac{4}{3}a_0^{-1}p_2\right)$, final expression for ψ_{40} is

$$\psi_{40}=\frac{a_0^{-3}}{4R^2\rho_{20}}\left[\beta_0^{-1}\left\{1-\frac{4}{3}m_0a_0^{-1}\left(\frac{R^2}{4s^2}-\frac{g_0^2}{R^2}-m_0\right)\right\}^2\right.$$
$$\left.+\frac{16}{9}m_0a_0^{-1}\left(\frac{R^2}{4s^2}-\frac{g_0^2}{R^2}-m_0\right)^2\right] \tag{86}$$

Analogous expressions for F_1', F_2' etc. can be obtained.

3. Symmetric properties of monochromator

Let us consiner the case in which $n=0$ and $\theta=0$, that is, $\phi_1=\sin\theta=0$ and $\phi_2=\cos\theta=1$. This case corresponds to the zero order spectrum or plane mirror-like reflection. In such a case, most terms of the optical path function L vanish. For such setting, consider the condition for the first order aberration to become negligible. First, the equation (70) becomes

$$F_1=\rho_{10}+\rho_{20}+\frac{1}{2}\psi_{11}w$$

and

$$F_1'=\rho_{10}'+\rho_{20}'+\frac{1}{2}\psi_{11}'w$$

When Fermat's principle is applied to (F_1+F_1'),

$$\psi_{11}+\psi_{11}'=0$$

Therefore, from (75) and (54)

$$\frac{y_0}{\rho_{20}}(1-a_0^{-1})+\frac{s_2}{\rho_{10}}a_0^{-1}+\frac{y_0'}{\rho_{20}'}(1-a_0'^{-1})+\frac{s_2'}{\rho_{10}'}a_0'^{-1}$$
$$-a_0^{-1}\left(s_2-\frac{a_0}{Rs}A_{11}\right)Q_0-a_0'^{-1}\left(s_2'-\frac{a_0'}{R's'}A_{11}'\right)Q_0'=0 \tag{87}$$

The conditions that satisfy the above equation are

$$\left.\begin{aligned} y_0' &= -\kappa y_0\,, & s_1' &= \kappa s_1\,, & s_2' &= -\kappa s_2\,, \\ z_1' &= -\kappa z_1\,, & R' &= \kappa R\,, & h_0 &= \kappa h_0'\,, \end{aligned}\right\} \tag{88}$$

and therefore, from $s(6)$, $g_0(8)$, $k_0(18)$, $m_0(23)$, $a_0(28)$, $\beta_0(34)$, $\rho_{10}(36)$, $\rho_{20}(36)$, $m_{11}(23')$, $A_{11}(37')$ and $Q_0(52)$

$$s' = \kappa s\,,\quad \rho_{10}' = \kappa\rho_{10}\,,\quad \rho_{20}' = \kappa\rho_{20}\,,\quad a_0' = a_0\,,\quad A_{11}' = -\kappa^3 A_{11}\,,\quad Q_0' = -\kappa^{-1}Q_0 \tag{88'}$$

where, $\kappa > 0$.

Since the case of $k \neq 1$ has no practical meaning, only the case of $\kappa = 1$ will be considered. Then, (88) is the conditions that M_1 and M_2-systems are symmetric. One of the optical arrangements satisfying the above conditions is shown in Fig. 3.

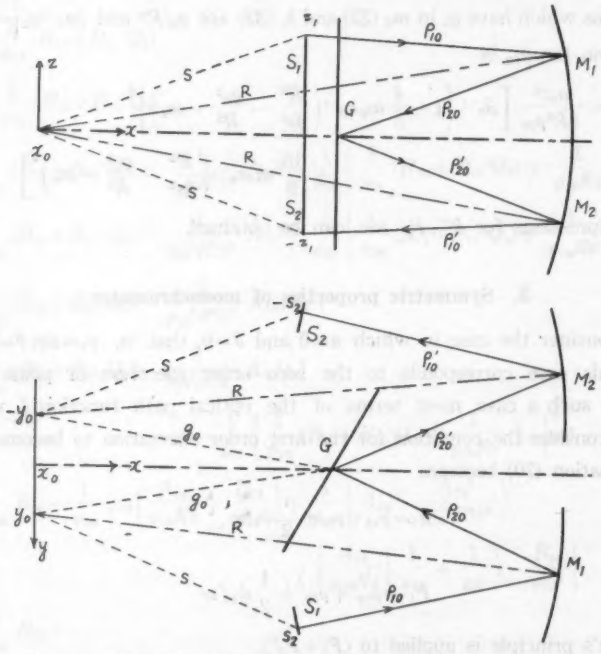


Fig. 3. Schematic diagram of symmetric optical arrangement of two mirrors system.

In such a case, even though $n \neq 0$ or $\theta \neq 0$, the conditions of (88) remain valid, and $\psi_{22}+\psi_{22}'$, $\psi_{31}+\psi_{31}'$, $\psi_{33}+\psi_{33}'$, $\psi_{12}z_1+\psi_{12}z_1'$ and $\psi_{39}z_1+\psi_{39}'z_1'$ all vanish because of their asymmetry. Therefore, it is sufficient to consider only the following terms instead of (70)~(73).

$$F_1+F_1'=2(\rho_{10}+\rho_{20})+\phi_1\psi_{10}w \tag{70'}$$

$$F_2+F_2'=(\phi_1^2\psi_{20}+\phi_2^2\psi_{21})w^2+(\phi_1^3\psi_{30}+\phi_1\phi_2^2\psi_{32})w^3 \tag{71'}$$

$$F_3+F_3'=\psi_{20}l^2 \tag{72'}$$

$$F_4+F_4'=\psi_{40}(w^2+l^2)^2 \tag{73'}$$

Although aberration expressions for Czerry-Turner type monochromator in general case are too complex to examine, if the system becomes symmetric, those expressions are simplified and we can easly derive various conditions for image formation.

4. Focal condition

Since the grating equation in the case of long slits can not be derived unless the focal condition is given, we shall consider it for the image formation in the first place. This should be determined from $\partial(F_2+F_2')/\partial w$. Therefore, from (71')

$$(\phi_1^2\psi_{20}+\phi_2^2\psi_{21})w+(\phi_1^3\psi_{30}+\phi_1\phi_2^2\psi_{32})w^2=0 \tag{89}$$

The solution that always satisfies the above equation independently of w and θ should be

$$\psi_{20}=\psi_{21}=\psi_{30}=\psi_{32}=0 \tag{90}$$

Inserting $A_1(37')$, $A_2(38')$, $B_1(54')$, $B_2(56')$ and $B_3(58')$ into (77), (78), (81) and (83), furthermore, using $m_1(23')$, $p_1(29)$ and $p_2(30)$, we see that, if $Q_0+\Delta=0$, (90) is satisfied, and since Δ is negligibly small, the focal condition is approximately given by

$$Q_0\equiv\frac{1}{\rho_{10}}-\frac{1}{\rho_{20}s}\left(Ra_0\beta_0^{\frac{1}{2}}-\frac{k_0^2}{s}\right)=0 \tag{52'}$$

This equation is very simple in form, but it is not so easy to obtain s or s_1 from it. However, approximate solution is obtained as follows.

Since

$$\frac{1}{\rho_{10}}+\frac{1}{\rho_{20}s}\left(Ra_0\beta_0^{\frac{1}{2}}-\frac{k_0^2}{s}\right)\neq 0$$

multiplication of (52') and the above expression gives

$$\varphi(s)\equiv\rho_{20}^2s^2-\rho_{10}^2\left(Ra_0\beta_0^{\frac{1}{2}}-\frac{k_0^2}{s}\right)^2=0 \tag{91}$$

Now, from (28), (34) and (23)

$$a_0\beta_0^{\frac{1}{2}}=1-\frac{1}{2}m_0+\frac{2}{3}p_0+\frac{1}{3}m_0p_0-\frac{1}{8}m_0^2+\cdots\cdots$$

$$=1-\frac{V_0^2}{2R^2s^2}+\frac{2}{3}\left(C_0+\frac{V_0}{R^2s^2}\right)D_0+\cdots\cdots \tag{92}$$

where

$$V_0 \equiv g_0^2 s^2 - k_0^2, \qquad C_0 \equiv 1 - \frac{R^2}{4s^2}, \qquad D_0 \equiv 1 - \frac{g_0^2}{R^2} \tag{93}$$

The values given in (28) and (34) make ρ_{10}^2 (36) as

$$\rho_{10}^2 = (R-s)^2 + \frac{V_0}{Rs}\left(1 - \frac{4}{3}p_0 + \frac{1}{4}m_0 + \frac{4}{3}p_0^2 + \cdots\cdots\right)$$

Similarly,

$$\rho_{20}^2 = \left(R - \frac{k_0^2}{s}\right)^2 - \frac{V_0}{Rs^2}\left(R - \frac{k_0^2}{s}\right)\left(1 - \frac{2}{3}p_0 + \frac{4}{9}p_0^2 + \cdots\cdots\right)$$
$$+ \frac{2V_0}{3s^2}p_0\left(1 - \frac{2}{3}p_0 + \cdots\cdots\right)$$

Therefore, the equation (91) is expressed as follows.

$$\varphi(s) = -\left\{R(R-2s) + \frac{V_0}{Rs}\right\}\left\{\left(R - \frac{k_0^2}{s}\right)^2 - \frac{V_0}{Rs^2}\left(R - \frac{k_0^2}{s}\right)\right\}$$
$$-\left\{(R-s)^2 + \frac{V_0}{Rs}\right\}\left[\frac{4}{3}RD_0\left(R - \frac{k_0^2}{s}\right)\left(C_0 + \frac{V_0}{R^2s^2}\right)\right.$$
$$\left. + \left\{\frac{V_0}{2Rs^2} - \frac{2}{3}RD_0\left(C_0 + \frac{V_0}{R^2s^2}\right)\right\}^2\right] + \cdots\cdots \tag{94}$$

Since $s \approx R/2$, approximate values of s and s_1 are

$$s = \frac{1}{2}R - \varphi\left(\frac{R}{2}\right)\left(\frac{\partial\varphi\left(\frac{R}{2}\right)}{\partial s}\right)^{-1} \tag{95}$$

$$s_1 = s\left(1 - \frac{r_0^2}{s^2}\right)^{\frac{1}{2}} \tag{96}$$

where,

$$r_0^2 \equiv s_2^2 + z_1^2$$

The following calculations are necessary to obtain the value of $\varphi(R/2)$

$$\left.\begin{aligned} &(s_1)_{s=R/2} = \frac{R}{2}\left(1 - \frac{4}{R^2}r_0^2\right)^{\frac{1}{2}} \\ &\left(R - \frac{k_0^2}{s}\right)_{s=R/2} = R\left(K_1 + \frac{2}{R^2}y_0 s_2\right) \\ &(V_0)_{s=R/2} = h_0^2 r_0^2 + Y_0 \end{aligned}\right\} \tag{97}$$

where,

$$\left.\begin{aligned} &K_1 \equiv 1 - \frac{h_0}{R}\left(1 - \frac{4}{R}r_0^2\right)^{\frac{1}{2}} \\ &Y_0 \equiv y_0\left\{\left(\frac{R^2}{4} - s_2^2\right)y_0 + Rh_0 s_2\left(1 - \frac{4}{R^2}r_0^2\right)^{\frac{1}{2}}\right\} \end{aligned}\right\} \tag{98}$$

On the assumption that h_0, the x-coordinate of the grating center, is near $R/2$ as in customary arrangement, $1-4D_0/3$ becomes negligibly small.

Hence

$$\varphi\left(\frac{R}{2}\right)=-2(h_0^2r_0^2+Y_0)\left\{\left(K_1+\frac{2}{R^2}y_0s_2\right)^2+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)\right\}$$
$$+8\frac{(h_0^2r_0^2+Y_0)^2}{R^4}\left(1-\frac{4}{3}D_0\right)\left\{K_1+\frac{2}{R^2}y_0s_2-\frac{1}{8}\left(1-\frac{4}{3}D_0\right)\right\}$$
$$-8\frac{(h_0^2r_0^2+Y_0)^3}{R^6}\left(1-\frac{4}{3}D_0\right)^2+\cdots\cdots$$
$$=-2(h_0^2r_0^2+Y_0)\left(K_1+\frac{2}{R^2}y_0s_2\right)^2\left\{1+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)\right\}+\cdots\cdots \quad (99)$$

Further, the following calculations are needed for obtaining $\partial\varphi(R/2)/\partial s$.

$$(s_1')_{s=R/2}\equiv\left(\frac{\partial s_1}{\partial s}\right)_{s=R/2}=\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}$$

from which

$$\left(R-\frac{k_0^2}{s}\right)'_{s=R/2}=-\frac{8h_0r_0^2}{R^3}\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}-\frac{4}{R^2}y_0s_2$$
$$(V_0')_{s=R/2}=y_0\left\{2h_0s_2\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}+Ry_0\right\}\equiv Y_1$$

whence we have

$$\frac{\partial\varphi(R/2)}{\partial s}=2R^3\left(K_1+\frac{2}{R^2}y_0s_2\right)^3+\frac{4(h_0^2r_0^2+Y_0)}{R}\left[\left(K_1+\frac{2}{R^2}y_0s_2\right)^2\right.$$
$$+\frac{4}{R^2}\left\{\frac{2h_0r_0^2}{R}\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}+y_0s_2\right\}\left(K_1+\frac{2}{R^2}y_0s_2+\frac{1}{3}D_0\right)$$
$$\left.+\frac{2}{3}D_0\left(1-\frac{4}{3}D_0\right)\right]-2Y_1\left[\left(K_1+\frac{2}{R^2}y_0s_2\right)^2\right.$$
$$\left.+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)\right]+\cdots\cdots$$
$$=2R^3\left(K_1+\frac{2}{R^2}y_0s_2\right)^3\left[1+\frac{2}{R^4}(h_0^2r_0^2+Y_0)\right.$$
$$\left.-\frac{Y_1}{R^3}\left\{1+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)^{-1}\right\}+\cdots\cdots\right] \quad (100)$$

which leads to

$$\left.\begin{aligned}s_C=&\frac{R}{2}+\frac{h_0^2r_0^2+Y_0}{R^3}\left\{1+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)^{-1}\right\}\left[1-\frac{2}{R^4}(h_0^2r_0^2+Y_0)\right.\\&\left.+\frac{Y_1}{R^3}\left\{1+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0s_2\right)^{-1}\right\}-\cdots\cdots\right]\\ s_{1C}=&(s_C^2-r_0^2)^{\frac{1}{2}}\end{aligned}\right\} \quad (101)$$

where, suffix C signifies the non-concentric *C*zerny-Turner type ($y_0 \neq 0$).

As for Ebert or Pfund type, since $y_0=0$ and therefore since $Y_0=Y_1=0$, we have, with suffix E to signify this type,

$$\left.\begin{aligned} s_E &= \frac{R}{2} + \frac{h_0^2 r_0^2}{R^3}\left(1+\frac{2}{3}D_0K_1^{-1}\right)\left\{1-\frac{2}{R^4}h_0^2r_0^2+\cdots\cdots\right\} \\ s_{1E} &= (s_E^2-r_0^2)^{\frac{1}{2}} \end{aligned}\right\} \qquad (102)$$

(101) and (102) lead to the conclusion that

(1) In the case of non-concentric Czerny-Turner type s_C and s_{1C} change as the coordinates (s_2z_1) of object point change, which means that, there is no fixed focal plane on which $x=$const.

(2) In the case of Ebert or Pfund type, if $r_0=$const., it becomes $s_E=$const. and $s_{1E}=$const. In other words, when object points are on the circle of radius $r_0(=y_1)$, image point are also on the same circle, symmetrically opposite to the former.

Since the grating center h_0 and the midpoint of slits s_{20} can be at any place unless the portion on the collimator surface to be used is designated, we can easily determine from (101) $s_C{}^0$ or $s_{1C}{}^0$ for the midpoint s_{20} of slit and from (102) s_E or s_{1E} for all image points as shown in Fig. 4. When the portion on the collimator surface to be used is designated, s_{20} is uniquely fixed if h_0 is given.

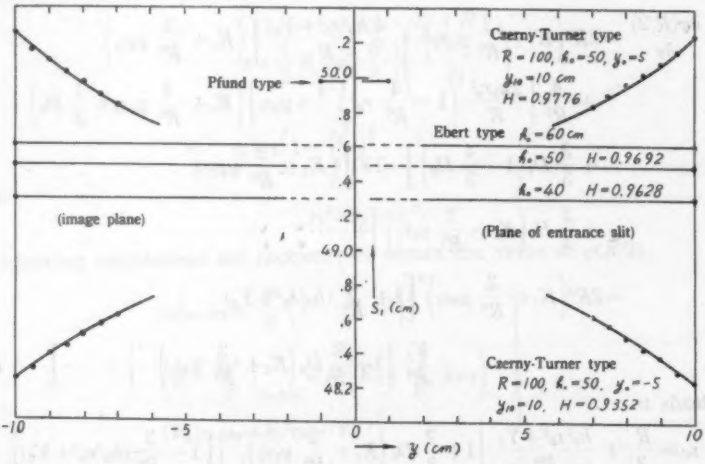


Fig. 4. Schematic diagram showing the relation of entrance slit and image plane in symmetric optical system.

But when in Ebert type $r_0 \neq$const.—the case of straight slit for example—focal points will not be on the plane on which $x=$const. In a symmetric system, the

image of straight slit is apt to be considered also straight in form. This may happen in the case of extremely short slit for example, but not necessarily so with a long slit because of the off-focus. In a monochromator, the form of slits must be determined to minimize the wavelength error arising from the turning of grating. For this determination, the grating equation should be discussed.

5. Grating equation and image form

Using (70′), (74), (54′) and focal condition (52′), we obtain for Fermat's principle

$$\frac{\partial}{\partial w}\left(F_1+F_1'+\frac{n\lambda}{d}w\right)=2\sin\theta\left\{\frac{s_1}{\rho_{10}}a_0^{-1}+\frac{h_0}{\rho_{20}}(a_0^{-1}-1)\right\}+\frac{n\lambda}{d}=0$$

Since the grating in Fig. 2 is used in negative order, let n be $-n$. Then, from (36) and (47), we have for the grating equation

$$n\lambda=2d\sin\theta\cdot H \tag{103}$$

where

$$H\equiv\frac{s_1}{s}a_0^{-1}\left\{1+\frac{R(R-2s\beta_0^{\frac{1}{2}})}{s^2}\right\}^{-\frac{1}{2}}$$
$$+\frac{h_0(a_0^{-1}-1)}{R-g_0}\left\{1+\frac{2R\left(g_0-\frac{k_0^2}{s}\beta_0^{\frac{1}{2}}-Rm_0a_0^{-1}\right)}{(R-g_0)^2}\right\}^{-\frac{1}{2}} \tag{104}$$

a) Case of Ebert or Pfund type.

We have seen that, in the case of Ebert or Pfund type, if s_1 and r_0 are given and therefore if s is given, the focal condition is satisfied for all the image points. Since $k_0^2=h_0s_1$, and both a_0 and p_0 are functions of only s, H is a function of s and s_1. Therefore, H becomes fixed. That is, the incident ray of wavelength λ from any point on a circular arc of radius $r_0(=y_1)$ with its center $y_0=0$ is imaged without wavelength error on the symmetrically opposite arc of the same circle, and the wavelength on the plane of image circle is changed only by the turning of grating. This image circle is none other than the well-known "Ebert circle" considered by Fastie. Therefore, Ebert type has a strong point; there is no wavelength error even if long slits are used.

The relation between object and image in Ebert type can be explained by Fig. 5. Let the grating G be a circle in form. In the figure, the outer circle C_1 is the collimating mirror, E is the slit circle or Ebert circle and C_2 is the intersecting circle of the collimating mirror with the sphere whose radius is ρ_{20} and center is coincident with the grating center. Since the circle C_2 is uniquely determined when Ebert circle is given, let it be named "Sub-Ebert circle." M and M' are portions on the

collimating mirror surface, the sizes of which correspond to that of the grating and their centers are on the sub-Ebert circle and, further, they are symmetric with respect to the grating center.

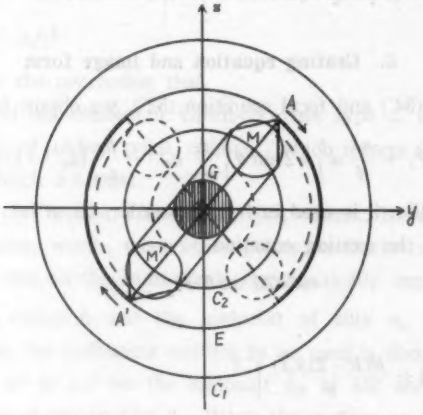


Fig. 5. Schematic diagram of image formation of Ebert type monochromator.
C_1, collimator; E, Ebert circle; C_2, sub-Ebert circle.

Now, the light from point A on the circle E is reflected to the grating G by the mirror portion M, and the diffracted ray is imaged on the point A' on the circle E by the mirror portion M'. If the object point A moves along Ebert circle E, M and M' move always along the sub-Ebert circle C_2 and the image point A' moves always along the circle E. Since A' and M' are always symmetric to A and M with respect to grating center, aberration by M, if any, is compensated. The figure shows also that the maximum movable range of M and M' along the "sub-Ebert circle" is the range in which the collimator surface is effectively used.

For Pfund type, because of y_1 being very small, the same as mentioned above is applicable provided that the slit is not long.

In the case of straight entrance slit, values of s, s_1, and k_0 vary along the slit. Hence H has various values despite the angle θ of the grating being fixed, which means that the wavelength of diffracted ray on the point symmetric to the object point differs from that of the image center. Thus, in order to minimize the wavelength error, form of the exit slit is bound to be made to differ from that of the entrance slit. Such a case is not a symmetric system in result. In such an asymmetric system, (60) that includes $\cos\theta$ should be looked into. Since the focal condition for a point on the image is different from that for any other point, a definite image

plane can not be fixed, and the inclusion of $\cot\theta$ in H can only be understood if it is on the assumption that every image point lies on the focal plane for the center of the image. Therefore, in the case of a straight entrance slit, the curve for the same wavelength, the "isochromat", will vary in form as θ is varied which fact will deprive the monochromator of its proper function. So we shall consider only for the symmetric system the conditions under which the entrance and exit slits of the same form can be used and H becomes a constant independent of the form of the slits.

b) Case of the non-concentric Czerny-Turner type

If y_0 is negligibly small, the case may be considered to be that of Ebert type. But if y_0 is too large to be neglected, all the conditions for image formation are affected by y_0. For the exit slit to give pure monochromatic ray, it should be placed along the isochromat. It is therefore convenient to have the form of the isochromat to be independent of the wavelength, for the slit form remains set. For this, H should be independent of the coordinates of the slits and remain as a constant which is H_0 for the midpoint $(s_{10}, s_{20}, 0)$ of the slit. As mentioned in the previous section concerning the focal condition, s_{10} is directly obtained from (101) when s_{20} is given. Thus, along the isochromat

$$H_0=H$$

From (101) and (104), H expressed by any point $(s_1\, s_2\, z_1)$ of the isochromat is a function of $s_2 z_1$. That is, it must be $H_0=H(s_2 z_1)$ Accordingly, put

$$\Phi(s_2 z_1)=0=H_0-H(s_2 z_1) \tag{105}$$

From (104) it is clear that predominant term in the grating equation is s_1/s and the effect of y_0 to the value of H is very small. Therefore, the difference between z_1 of non-concentric Czerny-Turner type and z_{1E} of Ebert type at the same s_2 will be very small, provided that h_0 and y_{10} are the same in both types. Thus z_1 becomes approximately

$$\begin{aligned} z_1 &= z_{1E}-\Phi(s_2 z_{1E})\Big/\frac{\partial\Phi(s_2 z_{1E})}{\partial z_1} \\ &= z_{1E}+\{H_0-H(s_2 z_{1E})\}/\{\partial H(s_2 z_{1E})/\partial z_1\} \end{aligned} \tag{106}$$

From (101), (23), (28), (34) and (98)

$$\left.\begin{aligned} \frac{\partial s}{\partial z_1} &\equiv s' = \frac{2h_0}{R^3} z_1\left[\left\{h_0-\frac{2}{R}s_2 y_0\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}\right\}\left\{1+\frac{2}{3}D_0\left(K_1+\frac{2}{R^2}y_0 s_2\right)^{-1}\right\}\right. \\ &\qquad \left.-\frac{4D_0}{3R^3}(h_0^2 r_0^2+Y_0)\left(K_1+\frac{2}{R^2}y_0 s_2\right)^{-2}\left(1-\frac{4}{R^2}r_0^2\right)^{-\frac{1}{2}}\right] \\ \frac{\partial s_1}{\partial z_1} &\equiv s_1' = (s^2-r_0^2)^{-\frac{1}{2}}(ss'-z_1) \end{aligned}\right\} \tag{107}$$

$$\left.\begin{aligned}
\frac{\partial m_0}{\partial z_1} &\equiv m_0' = -\frac{2s'}{R^3 s^3}(h_0^2 r_0^2 + Y_0) + \frac{2h_0 z_1}{R^2 s^2}\left\{h_0 - \frac{2}{R} s_2 y_0 \left(1 - \frac{4}{R^2} r_0^2\right)^{-\frac{1}{2}}\right\} \\
\frac{\partial a_0}{\partial z_1} &\equiv a_0' = \frac{2}{3} D_0\left(\frac{R^2}{2s^3} s' + m_0'\right) \\
\frac{\partial \beta_0}{\partial z_1} &\equiv \beta_0' = -a_0^{-2} m_0' + 2a_0^{-3} m_0 a_0'
\end{aligned}\right\}$$

Therefore,

$$\begin{aligned}
\frac{\partial H}{\partial z_1} = &\left\{\frac{a_0^{-1}}{s} s_1' - \frac{s_1}{s^2} a_0^{-1} s' - \frac{s_1}{s} a_0^{-2} a_0'\right\}\left\{1 + \frac{R}{s^2}(R - 2s\beta_0^{\frac{1}{2}})\right\}^{-\frac{1}{2}} \\
&+ \frac{s_1}{2s} a_0^{-1}\left\{1 + \frac{R}{s^2}(R - 2s\beta_0^{\frac{1}{2}})\right\}^{-\frac{3}{2}}\left\{\frac{2R}{s^3}(R - s\beta_0^{\frac{1}{2}}) s' + \frac{R}{s}\beta_0^{-\frac{1}{2}}\beta_0'\right\} \\
&- \frac{h_0}{R - g_0} a_0^{-2} a_0'\left\{1 + \frac{2R}{(R-g_0)^2}\left(g_0 - \frac{k_0^2}{s}\beta_0^{\frac{1}{2}} - R m_0 a_0^{-1}\right)\right\}^{-\frac{1}{2}} \\
&+ \frac{R h_0}{(R-g_0)^3}(a_0^{-1} - 1)\left\{1 + \frac{2R}{(R-g_0)^2}\left(g_0 - \frac{k_0^2}{s}\beta_0^{\frac{1}{2}} - R m_0 a_0^{-1}\right)\right\}^{-\frac{3}{2}} \\
&\left\{\frac{h_0}{s}\beta_0^{\frac{1}{2}} s_1' - \frac{k_0^2}{s^2}\beta_0^{\frac{1}{2}} s' + \frac{k_0^2}{2s}\beta_0^{-\frac{1}{2}}\beta_0' + R(a_0^{-1} m_0' - m_0 a_0^{-2} a_0')\right\} \qquad (108)
\end{aligned}$$

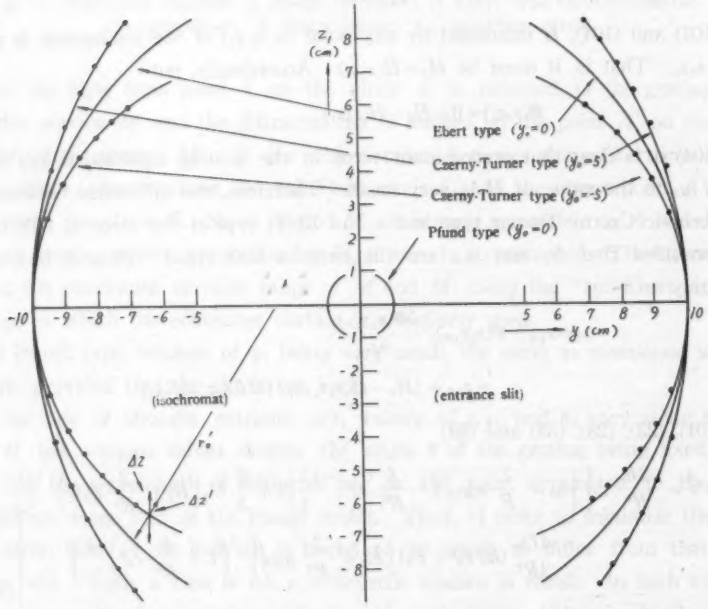


Fig. 6.

Inserting a given s_2 and also z_{1H} at s_2 into (101), (104), (107) and (103), we obtain z_1 from (106) and when those s_2 and z_1 are used, s or s_1 on isochromat is determined from (101). For example. results of numerical evaluation are shown in Figs. 4 and 6. From these figures, we see that, if $y_0<0$, the isochromat resembles an oblate ellipse and always $s_1 \geq s_{10}$, and if $y_0>0$, it resembles a prolate ellipse and always $s_1 \leq s_{10}$.

If, in Fig. 5, we replace the collimator mirror by two collimator mirrors, Ebert circle by entrance and exit semi-elliptic arcs and sub-Ebert circle by entrance and exit sub-semi-elliptic arcs, it will very well demonstrate how the image is formed in a non-concentric system. But this can not be tried in the case of long slits, for it is difficult to construct slits of such shape. Therefore, if for some reason, non-concentric Czerny-Turner type has to be used, y_0 must be made as small as possible, or else the slits must be short as possible as in the case of Littrow type.

6. Astigmatism

Astigmatism arises parallel to the grooves of grating and is small when the slit is short. Therefore, in the case of short slit, even if it is straight, the spectral resolution is not very much affected by astigmatism. However, as the slit becomes long, astigmatism is enhanced; therefore, even if the slit is of circular or elliptic arc, it must not be too long. For accurate determination of astigmatism, ray-tracing is needed, nevertheless we shall calculate it roughly as follows. Assuming that the focal plane is approximately perpendicular to ρ_{10}', we obtain approximately for the aberration Δz from the image point $z'(=-z)$ by the central ray as

$$\Delta z=\rho_{10}'\frac{\partial}{\partial l}(F_3+F_3') \tag{108'}$$

Since our optical system is symmetric, from (72')

$$\Delta z=H_2|l| \tag{109}$$

where,

$$H_2=2\rho_{10}\psi_{25} \tag{109'}$$

From ψ_{25}(80), B_{12}(54'), B_{25}(56'), A_{12}(37') and B_{12}(29), H_2 is given by

$$H_2/2=\frac{\rho_{10}}{\rho_{20}}(1-a_0^{-1})-z_1^2\left[\frac{\rho_{10}}{\rho_{20}}a_0^{-1}\left\{\frac{1}{\rho_{10}^2}\left(1+\frac{a_0}{Rs}A_{12}\right)^2-\frac{1}{s^2}\left(1-\frac{a_0}{Rs}A_{12}\right)^2\right\}\right.$$
$$\left.+\left(\frac{A_{12}}{\rho_{10}Rs}\right)^2-\frac{16\rho_{10}a_0^{-3}k_0^4}{9\rho_{20}R^2s^4}m_0D_0^2\right] \tag{111}$$

where,

$$A_{12}\equiv a_0^{-2}\beta_0^{-\frac{1}{2}}k_0^2\left(1-\frac{4}{3}m_0D_0\right)$$

Only the component $\Delta z'$ which is perpendicular to isochromat affects the spectral resolution, and this component is

$$\Delta z' = H_3 |l|$$

where,

$$H_3 \equiv H_2 \frac{z_1}{r_0'}$$

r_0' is the distance between the point $(s_2\, z_1)$ and the point at which the normal to the isochromat at the point $(s_2\, z_1)$ intersects y-axis (Fig. 6). The isochromats in the range of z_1 given in the figure are all approximately expressed by circular arcs, and from the figure we obtain $r_0' = 13.5$ cm for $y_0 = 5$ cm, and $r_0' = 7.5$ cm for $y_0 = -5$ cm. The result of numerical calculation of H_3 for the setting of Fig. 4 are shown in Fig. 7.

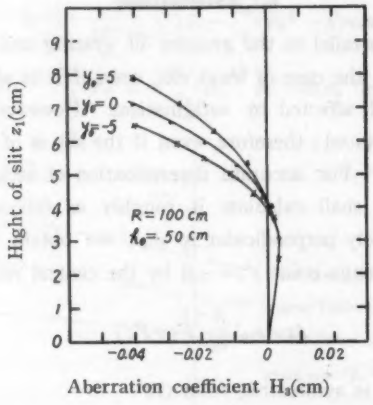


Fig. 7. Effect of astigmatism on spectral resolution.

From the figure, we see that the effect of astigmatism on the spectral resolution becomes gradually large as the slit length increases even if circular-arc-shaped slit (Ebert type) is used. Thus, even in Ebert type, slit length is restricted by astigmatism which can not be corrected. Furthermore, the figure shows that $\Delta z'$ increase in the order of the cases $y_0=5$, $y_0=0$, $y_0=-5$. Since Fig. 7 is the case of $y_{10}=$const., as seen from the values of H (Fig. 4), the portions of mirror surface to be used diverge from x-axis in the order of the cases $y_0=5$, 0, -5. That is, the degree of off-axis of the effective mirror surface with respect to the grating increases gradually. Therefore, such a large off-axial arrangement as $y_0=-5$ is not commendable.

7. Spherical aberration and optimum grating size

The quantity representing the spherical aberration is (F_4+F_4'). Therefore, in the symmetric system, from (73′)

$$F_4+F_4'=\psi_{40}(w^2+l^2)^2$$

ψ_{40} is given by the equation (86) and is independent of θ. Following Beutler, let us consider a circular grating of radius $\sigma=(w^2+l^2)^{\frac{1}{2}}$. The condition for image formation is

$$\psi_{40}\sigma^4 \leqq \frac{\lambda}{4}$$

Therefore,

$$\sigma \leqq \left(\frac{\lambda}{4\psi_{40}}\right)^{\frac{1}{4}}$$

Since the effect of y_0 is small, only the case of Ebert type will be calculated. In the case of the setting shown in Fig. 4, optimum diameter 2σ of the grating is

$$2\sigma \leqq 5.392\,\{\lambda(\mu)\}^{\frac{1}{4}}$$

2σ is shown in Fig. 8. If $h_0 \approx R/2$, $\rho_{20} \approx R/2$, ψ_{40} becomes approximately $(1/2R^3)$. Therefore,

$$\sigma \leqq \left(\frac{1}{2}R^3\lambda\right)^{\frac{1}{4}}$$

that is, the optimum grating size is approximately proportional to $R^{\frac{3}{4}}$ provided that the grating is ideal, and if not, the size is reduced.

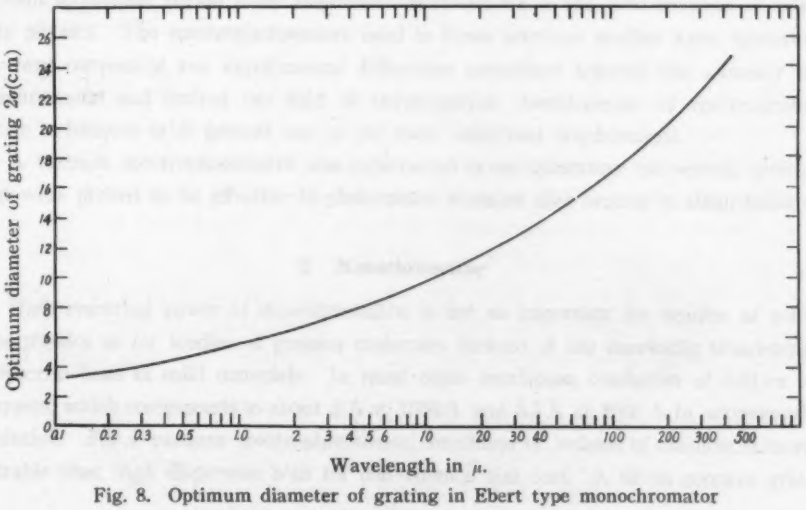


Fig. 8. Optimum diameter of grating in Ebert type monochromator in the case of $R=100$ cm, $h_0=50$ cm, $y_0=10$ cm.

8. Conclusion

From the foregoing, the following conclusion was drawn concerning the plane grating monochromator regarded as of perfect symmetric system.

a) With Ebert and concentric Czerny-Turner type monochromators provided with circular-arc-shaped slits, the wavelength error can be removed even if the slit are fairly long, but the slit length is restricted by astigmatism.

b) In the case of Czerny-Turner type in non-concentric setting, the form of isochromat is elliptic. Therefore, only the elliptic-arc-shaped slits can remove the wavelength error. However, when the coordinates of the centers of the spherical mirror change, the form of the ellipse also changes rendering the adjustment difficult.

c) In Pfund type, long slits are not used. Therefore, if there is no loss of light energy in bringing the ray onto the axis, this type may serve well in the near-infrared region, but not in far-infrared region.

d) Whichever the type it is, the effect of astigmatism is enhanced as the slit becomes long which rules out the use of long slits. However, this effect is minimized by setting the mirrors and the grating slightly off-axis.

e) The optimum grating size is proportional to $\lambda^{\frac{1}{2}}$ and approximately to $R^{\frac{1}{2}}$ in customary setup.

A Vacuum Spectrophotometer for Solid State Physics

Ryumyo ONAKA and Ikuo FUJITA

Institute for Optical Research, Tokyo University of Education
Shinjuku-ku, Tokyo

(Received May 10, 1960)

Abstract

A vacuum spectrophotometer with a 40 cm concave grating has been constructed for studies of solid state physics. A new method of mounting the grating developed by one of the authors is used in the monochromator in which the concave grating is rotated about an axis somewhat displaced from the center of the grating. With this arrangement, the resolution is found adequate for most purposes in the study of solid state physics. This spectrophotometer can be used for transmittance measurements in a wide wavelength range from 1050 A to 6000 A and in a temperature range down to the liquid nitrogen temperature. An X-ray irradiator is provided for coloration of crystals. Studies of fluorescence and phosphorescence are also possible.

1. Introduction

It has recently come into notice that studies of optical properties of solids in the vacuum ultraviolet region made important contributions to the advancement of solid state physics. The spectrophotometers used in those previous studies were, however, not very convenient and experimental difficulties sometimes lowered the accuracy of measurements and limited the field of investigation. Development of spectrophotometric techniques is at present one of the most important requirements.

A vacuum spectrophotometer was constructed in our laboratory and several devices used were proved to be effective in photometric work on color centers in alkali halides.

2. Monochromator

High resolving power of monochromator is not so important for studies of solid state physics as for studies of gasious molecules because of the inevitable broadening of spectral lines in solid materials. In most cases maximum resolution of 0.01 ev is adequate, which corresponds to about 2 A at 1500 A and 3.5 A at 2000 A in wavelength resolution. For a vacuum spectrophotometer, smallness of volume of chamber is more desirable than high dispersion both for convenience and cost. A 40 cm concave grat-

ing with 600 grooves per mm was used in the present spectrophotometer, and the small spectrophotometer that was built was found capable in separating two lines 3 A apart. This limit of resolving power is not due to the smallness in size of the monochromator, but to insufficient intensity of the light source used.

Several types of grating mounting for vacuum monochromators have already been developed. As to the mounting condition for the present work, the following are desirable to be fulfilled.

1. *Fixed light beam.* The direction of the light beam emerging from the exit slit is to be fixed, otherwise the light beam would illuminate different points on the specimen when the wavelengths are changed. The mounting of a concave grating on the end of an arm pivoted at the center of the Rowland circle used by Fujioka and Ito,[1] and by Tousey *et al*,[2] is, therefore, not suitable for the present purpose in spite of its excellent focusing. Rotation of a concave grating around an axis at or near the center of the grating is desirable.

2. *Good efficiency.* Efficiency in energy is important, for the reflective power of the grating surface is very low in the far ultraviolet region. Except the case of extreme grazing incidence, normal incidence is desirable, for the astigmatism is small and almost full energy emerges through the exit slit.

3. *Little polarization at reflection by the grating.* For this, normal incidence is again preferred.

4. *Sharp focusing.* Sharp focusing becomes important in the short wavelength range, for the required resolution is necessarily high. Resolution of 0.01 ev corresponds to 0.8 A at 1000 A, but to 20 A at 5000 A. In the long wavelength range it matters little even if the focusing is not sharp.

In the present spectrophotometer, the concave grating is rotated about a vertical axis displaced a little from the center of the grating. Blurring of image due to rotation is reduced by simultaneous translation of the grating. This mounting of grating was invented by Johnson[3] and later improved by one of the present authors.[4] The monochromator is shown schematically in Fig. 1. The entrance and exit slits S_1, S_2 are fixed in the directions making an angle of about 30 degrees at the concave grating G. The axis of rotation of the grating is at O, displaced on the plane of rotation perpendicularly to bisecter of the angle S_1GS_2. The optimum displacement, with which the monochromator can maintain good focus as the grating is rotated, is given by the

1. Fujioka and Ito, Science of Light, **1** (1951) 1.
2. Tousey, Johnson, Richardson and Toran, J. Opt. Soc. Am., **41** (1951) 696.
3. P.D. Johnson, Rev. Sci. Inst., **28** (1957) 833.
4. Ryumyo Onaka, Science of Light, **7** (1958) 23.

following equation

$$l = R \sin\theta / \{1 - \tan\phi\,(\tan\alpha_0 - \tan\beta_0)/2\}\ ,$$

where l: optimum displacement of the center of rotation,
θ: angle between the grating normal and the bisecter of segment $S_1 S_2$,
ϕ: half of the angle S_1GS_2,
α_0: angle of incidence
β_0: angle of diffraction
} when the monochromator is set for the wavelength taken as standard.

These quantities are related one another as

$$\alpha_0 = \theta + \phi\ ,$$
$$\beta_0 = \theta - \phi\ .$$

The sine-bar scanning mechanism A (Fig. 1) turns the shaft which is directly connected through a vacuum-tight sleeve to the turn table for the grating. The wavelength of emerging light from the monochromator is, therefore, linearly proportional

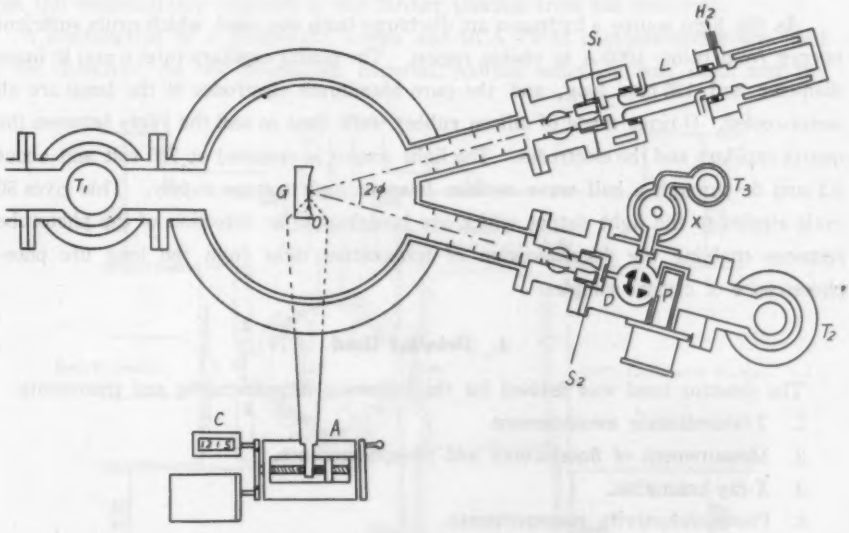


Fig. 1. Vacuum Spectrophotometer

G: Grating; O: Center of Rotation; A: Sine-bar Scanning Mechanism; C: Wavelength Counter; M: Synchronous Motor S_1: Entrance Slit; S_2: Exit Slit; L: Light Source; D: Detector Head; P: Phototube Housing; X: X-ray Tube; T_1, T_2, T_3: Cold Traps and Diffusion Pump.

to the advance of the scanning screw and is indicated on a digital counter in Angstrom units.

The entrance and exit slits are both of symmetrically opening type, the width of which are made adjustable from outside of the vacuum chamber. Between the slits and the monochromator body, large stop cocks are provided to facilitate disjoining of the light source or the detector head from the vacuum chamber whenever needed.

Three vacuum systems are incorporated, one each for the monochromator, the detector head and the additional equipment consisting of an X-ray tube, an evaporation chamber etc. The vacuum system for the monochromator comprises a liquid nitrogen trap, a 3 inch oil diffusion pump and a rotary oil pump.

For measurements at wavelengths longer than 1050 A, a lithium fluoride window is provided to isolate the detector head without impairing its high vacuum. This allows the monochromator chamber to be filled with pure hydrogen at appropriate pressure for the operation of the light source. The hydrogen gas seems to protect the grating against sputtering of the light source.

3. Light Source

As the light source a hydrogen arc discharge lamp was used, which emits sufficient energy from below 1000 A to visible region. The quartz capillary tube, 6 mm in inner diameter and 120 mm long, and the pure aluminium electrodes of the lamp are all water-cooled. O-rings made of silicon rubber were used to seal the joints between the quartz capillary and the electrodes. The light source is operated at 700 volt and about 0.3 amp dc through a half wave rectifier from ac high voltage supply. This gives 50 cycle ripples to the light output which are beneficial to ac detection of the phototube response enabling the discrimination of transmitted light from the long life phosphorescence of crystal samples.

4. Detector Head

The detector head was devised for the following measurements and treatments.

1. Transmittance measurement.
2. Measurement of fluorescence and phosphorescence.
3. X-ray coloration.
4. Photoconductivity measurements.
5. Sample making by vacuum evaporation.

(4 and 5 are in prospect. The present X-ray tube is intended to be replaced by an evaporation equipment when needed.)

The detector head is shown in Fig. 2. For transmittance measurements, the specimen crystal held in a holder is placed in contact with the bottom of a cryostat

and can be cooled down to the liquid nitrogen temperature. The position of the crystal is about 7 mm away from the axis of the cryostat cylinder, which allows the crystal to be in or out of the light beam by rotating the cryostat. The cryostat is sealed vacuum tight by means of an O-ring. Atmospheric pressure on the cryostat is borne by a thrust bearing.

Two positions are possible for the crystal to be in the light beam, near to and far from the light detector. The crystal faces the light detector (fluorescent screen) with different solid angle at these two positions, and scattered light at the crystal surface and fluorescence of the crystal, if any, give, therefore, two different contributions to the response of the detector. The radio between these two components in the response depends only upon the geometrical relation between the detector and the crystal. Thus it is possible to find this excess of response from two readings of the detector. The ratio of the excess component by one reading to that by the other was approximately one to four both in theoretical estimation and experimental determination by the use of a fluorescent crystal. The true transmittance is therefore found if a third of the difference between the two readings obtained at above two positions is substracted from the transmittance obtained at the farther position from the detector.

A combination of a fluorescent screen and RCA 1 P 28 photomultiplier was used as the detector. As the fluorescent material, sodium salicylate was used, and the

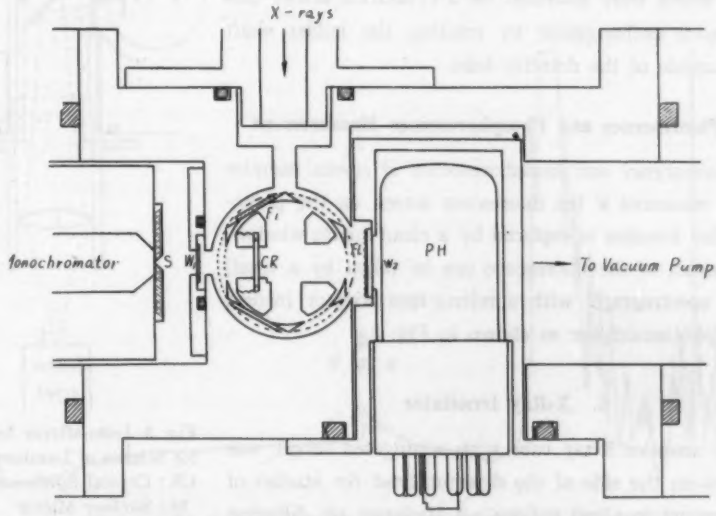


Fig. 2. Detector Head. S: Exit Slit, W_1: Li F Window, CR: Crystal Specimen, F_1: KBr-Filter, F_2: Glass-Filter, FL: Fluorescent Material, W_2: Quartz Window, PH: Photomultiplier 1P28

quartz window of the photomultiplier housing was coated with it.

The output of the photomultipler is fed to a dc or ac amplifier and then to an electronic recorder. When the sample crystals had phosphorescence as the result of X-ray irradiation for example, the ac amplifier tuned to 50 cycle was used. Since the light source had 50 cycle ripples, 50 cycle component of the transmitted light enabled the true transmittance of the specimen to be obtained, for the afterglow becomes undetectable when an ac amplifier is used.

When the scattered light in the monochromator is negligible as in the case of wavelengths longer than 1350 A, addition of a logarithmic amplifier facilitates the calculation of transmittance.

With a 2 inch oil diffusion pump and a liquid nitrogen trap, pressure of the order of 10^{-7} mm Hg is attained in the detector head.

When the vacuum spectrophotometer is used for near ultraviolet or visible region, use of a set of filters becomes necessary to separate higher order spectra. The following filters were used.

Up to 1050 A	No filter	No window
from 1050 A to 2100 A	No filter	LiF window
from 2100 A to 3500 A	KBr filter	LiF window
from 3500 A to 6000 A	Glass filter	LiF window

These filters were mounted on a cylindrical holder and were made exchangeable by rotating the holder shaft from outside of the detector head.

4. Fluorescence and Phosphorescence Measurement

Fluorescence and phosphorescence of crystal samples can be measured if the fluorescent screen on the photomultiplier housing is replaced by a clean quartz window. The spectra of the fluorescene can be taken by a small quartz spectrograph with a mirror-lens adaptor instead of the photomultiplier as shown in Fig. 3.

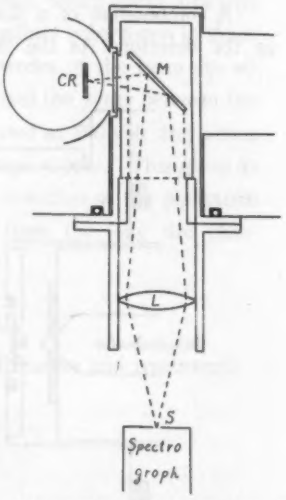


Fig. 3. Lens-Mirror System for Studies of Luminescenc.
CR: Crystal Specimen
M: Surface Mirror
L: Quartz Lens
S: Entrance Slit of Spectrograph.

5. X-Ray Irradiator

An unsealed X-ray tube with a tungsten target was provided on the side of the detector head for studies of color centers in alkali halides. A Hickman oil diffusion pump was used for evacuation. The X-ray tube is operated at 30 KV 25 mA in most cases.

6. Performance

The spectrophotometer can be used over a wide wavelength range from 1050 A to 6000 A with a lithium fluoride window and a hydrogen discharge tube. Wavelength accuracy was within ± 2 A in error throughout the above wavelength range, and the best resolution obtained was about 3 A at the vacuum region (below 2000 A).

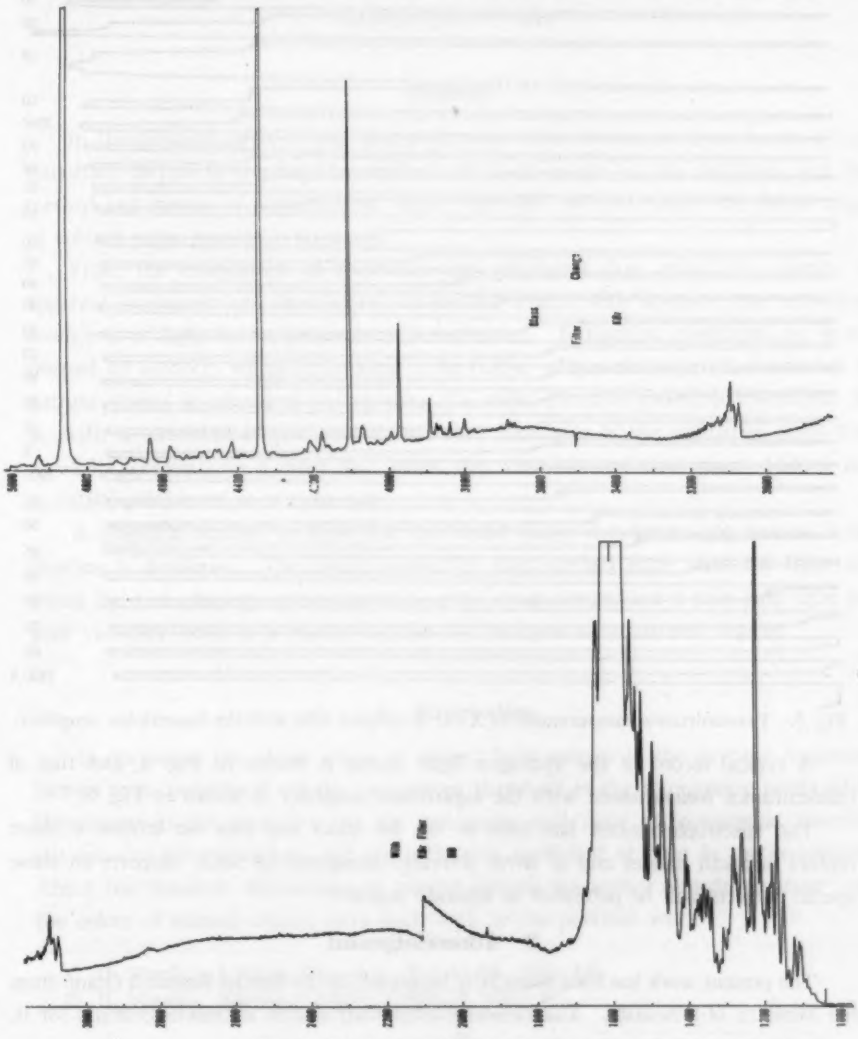


Fig. 4. Spectrum of the hydrogen arc.

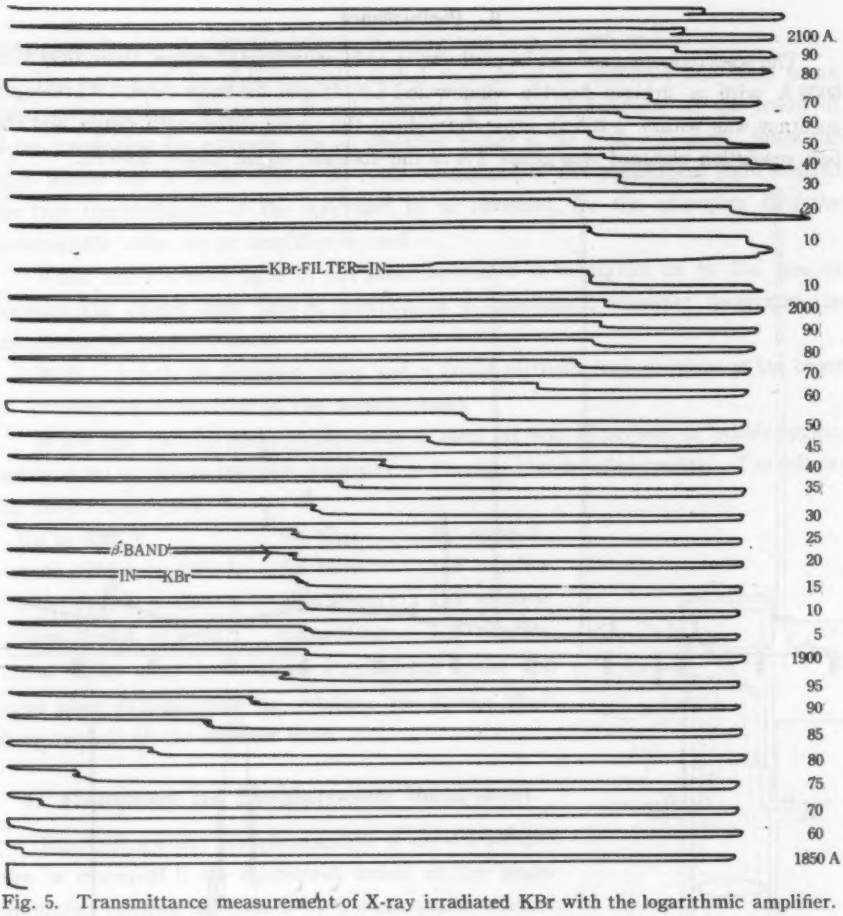


Fig. 5. Transmittance measurement of X-ray irradiated KBr with the logarithmic amplifier.

A typical record of the hydrogen light source is shown in Fig. 4, and that of transmittance measurement with the logarithmic amplifier is shown in Fig. 5.

The spectrophotometer has been in use for about one year for studies of color centers in alkali halides and of silver activated phosphors of NaCl. Reports on these special subjects will be published in separate papers.[5]

7. Acknowledgement

The present work has been financially supported by the Special Research Grant from the Ministry of Education. The authors express their sincere acknowledgements for it.

5. Onaka and Fujita, Phys. Rev., to be published.

Apparent Colors of Natural Objects (II)

Hikaru MASAKI

Railway Technical Research Institute, Kokubunji-machi, Tokyo

(Received June 10, 1960)

Summary

In the previous paper it was stated that the color of natural object is one of the important factors in treating the problem of visual range in the daytime, and the method and results of measurement were reported. In this paper the visual range of colored point sources is discussed.

From the comparison of measured and calculated data about the change in apparent luminance and chromaticity of natural objects with distance, the extinction coefficient of light in the atmosphere is estimated. Extinction coefficient, σ_λ, is expressed by $\sigma_0(\lambda_0/\lambda)^n$, where λ_0 is taken to be $0.55\,\mu$. Then the apparent luminance of natural objects is calculated by replacing the mean effective extinction coefficient by σ_0 with a precision of two percent in error. Examples of the estimation show that $\sigma_\lambda=2.5(\lambda_0/\lambda)^{0.8}\,\mathrm{km}^{-1}$ on a rainy day when the visual range was about 1.5 km, and $\sigma_\lambda=0.04(\lambda_0/\lambda)^{2.7}\,\mathrm{km}^{-1}$ on a clear day.

A graphical method to determine the visual range of colored light sources in the daytime is developed. The results gained by this method show that red light and green light of adequate saturation have good visual ranges but a pale blue light has poor visibility when it is viewed against the background of natural objects.

1. Introduction

In discussing the visual range of colored light sources in the daytime, important factors to be considered are the recognition threshold of the illuminance produced by the sources at the observer's eye, the luminance and color of background, therefore the colors of natural objects, and the extinction coefficient of light in the atmosphere. About the threshold illuminance of colored sources the author already reported[1], and the colors of natural objects were dealt with in the previous work[2].

1) H. Masaki: J. Illum. Eng. Inst. (Tokyo), **39**, (1955) 430.
2) H. Masaki: Science of Light (Tokyo), **8**, (1959) 67.

The extinction coefficient σ_λ is the most difficult factor to obtain experimentally. It is known that the change in apparent colors of natural objects with distance relates to σ_λ. If this relation can be used for the estimation of σ_λ, it will be a great advantage in treating the problem of visibility.

A graphical method to estimate σ_λ by simply measuring the apparent colors of objects at various distances was introduced and applied in computing the visual range.

2. Change in colors of natural objects with distance

The apparent colors of natural objects change from yellow-green or yellow to blue with increasing distance and finally approach the color of the horizon sky. The luminances also approach that of the sky. This is because the light from objects attenuates in the atmosphere and the daylight scattered in the path is added to it. The equation expressing the apparent luminance B_r of the object observed at the distance r was first derived by Koschmieder and arranged to be more comprehensive by Duntley[3] and Middleton[4, 5]. It is

$$B_r = B_h(1 - e^{-\sigma r}) + B_0 e^{-\sigma r}, \qquad (1)$$

where B_h is the luminance of the horizon sky, B_0 the luminance of the object observed at close distance, and σ the extinction coefficient. The first term of the right side expresses the contribution of the scattered daylight to the luminance and the second term the contribution of the attenuated light from the object. Applying the equation (1) to a monochromatic light of wavelength λ, we obtain the following equation

$$B_{r,\lambda} = B_{h,\lambda}(1 - e^{-\sigma_\lambda r}) + B_{0,\lambda} e^{-\sigma_\lambda r}. \qquad (2)$$

Middleton calculated the change of colors by integrating the equation (2) over the wavelength. He assumed that the light from the horizon sky had the same spectral distribution as the prevailing illumination and expressed $B_{0,\lambda}$ in the form

$$B_{0,\lambda} = f_\lambda (C_w + 1) B_{h,\lambda} \qquad (3)$$

where f_λ is the spectral reflectance of the object and C_w the contrast between the luminance of an ideal white object in the vicinity of the observer and that of the horizon sky. As for the spectral intensity distribution of the sky, that of the CIE source C was used, which represents well the general atmospheric conditions. But when the prevailing illumination and the light from the horizon sky have different spectral qualities, for instance when on a clear day the object is illuminated by the

3) S. Q. Duntley: J. Opt. Soc. Amer. **38**, (1948) 179.
4) W.E.K. Middleton: J. Opt. Soc. Amer. **40**, (1950) 373.
5) W.E.K. Middleton: Vision Through the Atmosphere, Toronto Univ. Press, 1952.

direct sunlight of the color temperature 5,500°K and the light from the horizon sky has the color temperature of more than 10,000°K, the following more precise approximation is obtained by integrating (2) directly:

$$X_r=\int_V^R \{B_{h,\lambda}(1-e^{-\sigma_\lambda r})+B_{0,\lambda}e^{-\sigma_\lambda r}\}\bar{x}_\lambda d\lambda\,, \quad \text{etc.} \tag{4}$$

In Fig. 1 the changes in the chromaticity of foliage ($x_0=0.3292$, $y_0=0.3778$) against the blue sky ($x_h=0.2750$, $y_h=0.2954$) and the sky of the same spectral intensity

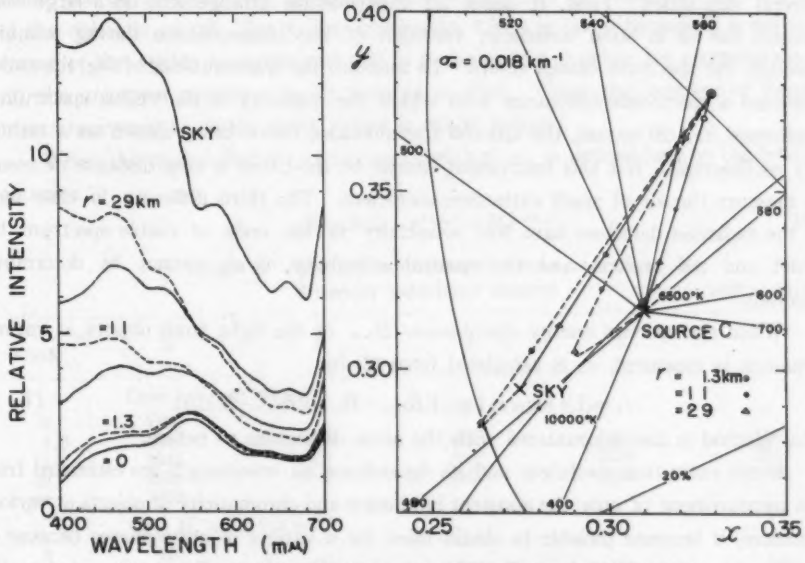


Fig. 1. Spectral distributions and chromaticity coordinates of the light from foliage at various distances in non-selective (full lines) and moderately selective (dashed lines) atmospheres (calculated results).

distribution as the source *C*, calculated by (4) for non-selective and moderately selective air, are shown. The change in the spectral distribution by distance is also shown. It might not be reasonable to consider the sky to be blue despite the atmosphere being non-selective, but it is only shown as an example of calculation. The calculated results show that the change is greater against the blue sky than against the sky with the characteristic of the source *C*. It is worth noting that the original color of the object is lost at a small distance even in a clear atmosphere.

Use of an electronic computor saved the otherwise laborious work of numerical integration of (4).

3. Estimation of extinction coefficient

The general method to obtain the extinction coefficient σ of light in the atmosphere is first to measure the transmittance τ_r through the layer of air of the thickness r, then compute by the equation

$$\tau_r = e^{-\sigma r}. \tag{5}$$

By measuring the transmittance spectrophotometrically, the extinction coefficient σ_λ of light for the wavelength λ is obtained[6,7]. But this method is attended with several difficulties. First, it needs an experimental arrangement on a large scale. Second, the air is often unsteady; variation of the transmittance during scanning through the spectrum causes errors. To measure the transmittance of fog, the author designed a spectrotelephotometer with which the scanning of the visible spectrum is performed in 1/20 second, the spectral transmittance curve being shown on a cathode ray oscillograph. But this instrument cannot be used over a long distance necessary to measure the air of small extinction coefficient. The third difficulty is that most of the radiation detectors have low sensitivity in the ends of visible spectrum, the violet and red regions, and the spectral selectivity of σ_λ cannot be determined precisely.

When the spectral energy distribution $B_{r,\lambda}$ of the light from objects at various distances is measured, σ_λ is calculated from (2) by

$$\sigma_\lambda = (1/r \log_{10} e) \log_{10} \{(B_{h,\lambda} - B_{0,\lambda})/(B_{h,\lambda} - B_{r,\lambda})\}. \tag{6}$$

This method is also encountered with the same difficulties as before.

If the extinction coefficient and its dependence on wavelength are estimated from the measurement of only the apparent luminance and chromaticity of objects at various distances, it becomes possible to obtain them for a variety of atmospheres because of the experimental simplicity. For this purpose, equations (4) are converted to the summation in the form

$$X_r = \sum_i \{B'_{hi} + (B_{0i} - B_{hi}) e^{-\sigma_i r}\} \bar{x}_i, \quad \text{etc.} \tag{7}$$

When B_{hi} and B_{0i} are known and X_r, Y_r and Z_r are measured at more than $i/3$ distances, σ_i is determined by solving the simultaneous equations (7) numbering i which is unknown at present. It is not, however, easy to solve such equations even by an electronic computor. Besides, the measured values of X_r, Y_r and Z_r may contain experimental errors, hence the mean values of σ_i are to be determined by numbers of measurements as many as several times of $i/3$.

Fortunately it is known[5] that σ_λ or σ_i is approximately inversely proportional to

6) A. Arnulf, J. Bricard, É. Curé et C. Veret: Rev. Opt. 38 (1959) 105.
7) H. Masaki: Railw. Tech. Res. Inst. Interm. Rep. 7-120, (1956).

a power of the wavelength λ and the exponent n lies in the range of about 0—4. σ_i is then expressed as

$$\sigma_i = \sigma_0 (\lambda_0/\lambda_i)^n = k_{ni}\sigma_0 . \tag{8}$$

Applying (8) to (7) we obtain

$$\left.\begin{aligned} B_{ri} &= \{B_{hi}+(B_{0i}-B_{hi})e^{-k_{ni}\sigma_0 r} \\ X_r &= \sum_i B_{ri}\bar{x}_i,\ Y_r = \sum_i B_{ri}\bar{y}_i,\ Z_r = \sum_i B_{ri}\bar{z}_i \end{aligned}\right\}, \tag{9}$$

where the units of B_{hi} and B_{ri} are so determined that Y_r expresses the luminance of the object in cd/m². For λ_0 the wavelength 0.55 μ is taken because it is nearly the center of the visible spectrum and 0.55 is a convenient number for computation. In (9), the unknown number i in (7) reduces to two. Thus the number of measurements necessary to obtain mean values is much reduced.

When the mean effective extinction coefficient σ_m is defined here by the following equation

$$B_r = B_h(1-e^{-\sigma_m r})+B_0 e^{-\sigma_m r} \tag{10}$$

Table 1. Apparent luminance of foliage at various distances (calculated results). unit: cd/m²

r (km) \ n	0	1	2	3	4
	Case I, $\sigma_0=0.040\ \mathrm{km}^{-1}$				
0	1150				
1.3	1349	1350	1352	1356	1361
11	2545	2551	2560	2576	2589
29	3839	3844	3847	3851	3846
∞	5050				
	Case II, $\sigma_0=0.045\ \mathrm{km}^{-1}$				
0	420				
0.3	559	559	560	561	563
5	2527	2522	2525	2542	2556
21	6814	6796	6778	6772	6745
36	8809	8787	8752	8717	8653
∞	10890				
	Case III, $\sigma_0=0.200\ \mathrm{km}^{-1}$				
0	240				
0.3	353	353	353	354	355
0.75	512	511	511	513	515
3.5	1222	1218	1215	1216	1213
5.0	1473	1468	1464	1461	1454
∞	2190				
	Case IV, $\sigma_0=2.5\ \mathrm{km}^{-1}$				
0	130				
0.05	198.2	197.8	197.8	198.3	198.7
0.3	436.3	434.9	433.6	433.3	432.0
0.75	621.6	619.8	617.3	614.5	610.0
1	663.0	661.5	659.0	656.1	651.5
∞	711				

it is found from the result of more than hundred calculations, some of which are shown later (Table 1), that the error in calculated values of B_r caused by replacing σ_m by σ_0 in (8), the extinction coefficient for the wavelength 0.55 μ, is less than two percent. This happens because the spectral distribution of the light from natural objects is not conspicuously selective. By putting $\sigma_m=\sigma_0$, σ_0 is obtained only from the change in luminance with the distance, namely

$$\sigma_0=(1/r\log_{10}e)\log_{10}\{(B_h-B_0)/(B_h-B_r)\} \tag{11}$$

It is more easily estimated graphically as shown in Examples.

The exponent n in (8) represents the degree of dependence of σ_λ or σ_t on wavelength. As it is considered to be about $0-4$ in the ordinary atmosphere, chromaticity coordinates of the light from objects at various distances are calculated by (9) with σ_0 determined by (11) and $n=0$, 1, 2, 3, and 4. It needed somewhat laborious work to calculate manually. But once a program for the electronic computor was formulated, many computations could be made in a short time. Then by comparing the measured data with the calculated on the chromaticity diagram, n is easily determined. Examples are shown in Table 1 and Figs 2~5.

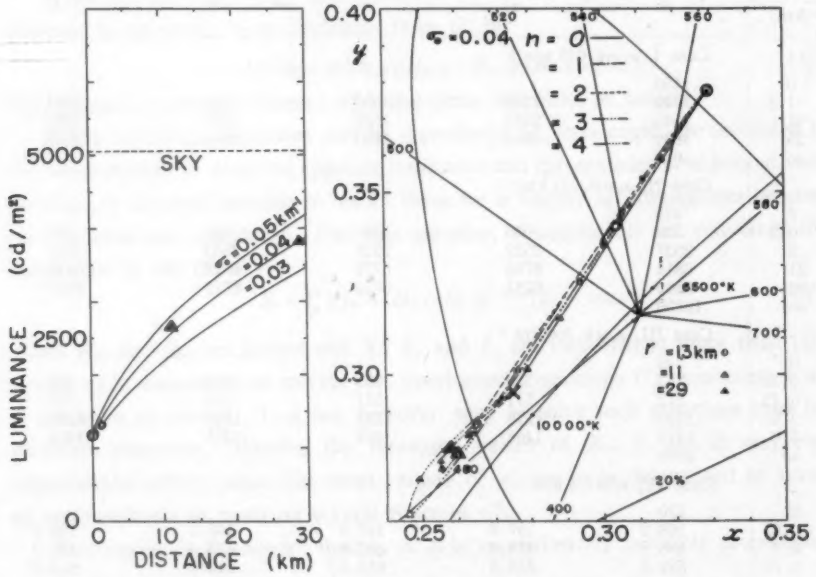


Fig. 2. Luminances and chromaticities of foliage at various distances (case I). Symbols with heavy outlines represent experimental results.

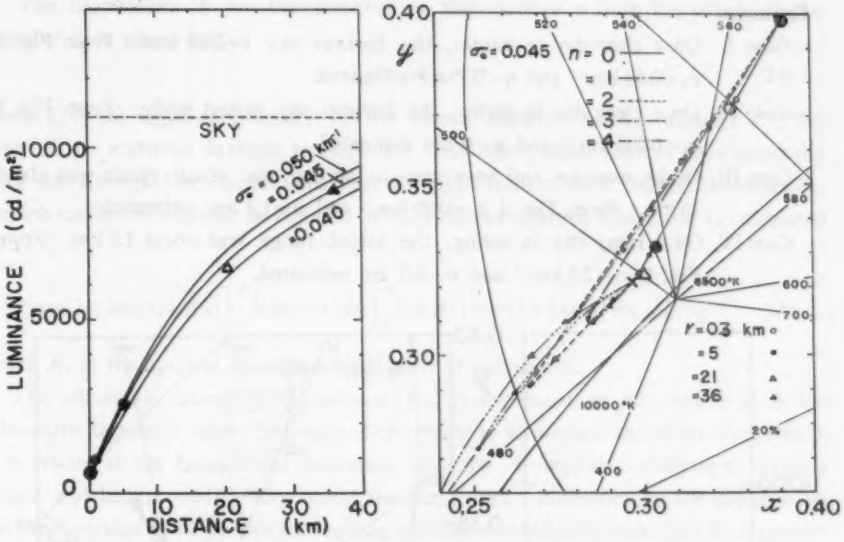


Fig. 3. Luminances and chromaticities of foliage at various distances (case II). Symbols with heavy outlines represent experimental results.

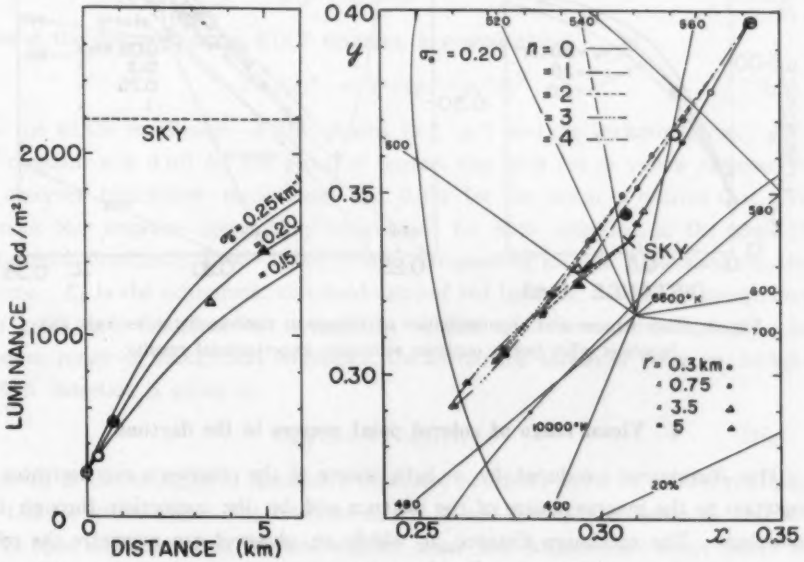


Fig. 4. Luminances and chromaticities of foliage at various distances (case III). Symbols with heavy outlines represent experimental results.

Examples

Gase I. On a clear day in winter, the horizon sky looked blue. From Fig. 2 $\sigma_0=0.04\,km^{-1}$ and $n=2.7$ are estimated.

Case II. On a clear day in spring, the horizon sky looked white. From Fig. 3 $\sigma_0=0.045\,km^{-1}$ and $n=0$ are estimated.

Case III. On an overcast and hazy day in spring, the visual range was about 10 km. From Fig. 4 $\sigma_0=0.20\,km^{-1}$ and $n=1.3$ are estimated.

Case IV. On a rainy day in spring, the visual range was about 1.5 km. From Fig. 5 $\sigma_0=2.5\,km^{-1}$ and $n=0.5$ are estimated.

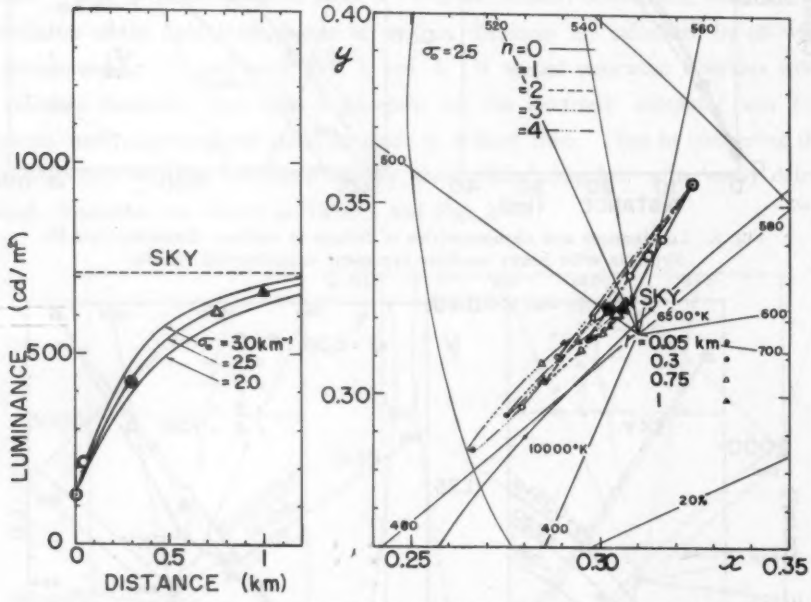


Fig. 5. Luminances and chromaticities of foliage at various distances (case IV). Symbols with heavy outlines represent experimental results.

4. Visual range of colored point sources in the daytime

The illuminance produced by a light source at the observer's eye decreases in proportion to the inverse square of the distance and by the extinction through the atmosphere. The maximum distance at which an observer can recognize the color of the source is the distance at which the illuminance decreases to its chromatic threshold value.

The illuminance E for the observer at the distance r from the source of the intensity I_0 is given by

$$E = I_0 e^{-\sigma r}/r^2 = I_r/r^2 \tag{12}$$

where I_r is the apparent intensity of the source observed at the distance r if the effect of the scattered daylight is neglected. As the colored sources have generally more selective spectral distributions than the light from natural objects, the mean extinction coefficient σ cannot be replaced by σ_0 of the wavelength 0.55 μ. It is calculated by

$$\sigma = -(1/r \log_{10} e) \log_{10}\left(\int_V^R E_\lambda \bar{y}_\lambda e^{-\sigma_\lambda r} d\lambda \Big/ \int_V^R E_\lambda \bar{y}_\lambda d\lambda\right) = -(1/r \log_{10} e) \log_{10}(I_r/I_0) \tag{13}$$

where E_λ is the spectral intensity distribution of the source.

The chromatic threshold illuminance $E_{t\sigma}$ is expressed as pE_t, where E_t is the achromatic threshold value. The ratio of chromatic to achromatic threshold illuminance, p, is related to the background luminance and the chromaticity difference between source and background[1]. Expressing the chromaticity difference by the distance on the Breckenridge and Schaub's rectangular uniform chromaticity scale (RUCS) diagram[8] the author determined the relation as

$$d(p-1)^{\frac{1}{2}} = a \tag{14}$$

where d, the distance on the RUCS diagram, is computed by

$$d = [(x_s'' - x_b'')^2 + (y_s'' - y_b'')^2]^{\frac{1}{2}} \tag{15}$$

from the RUCS coordinates of the source (x_s'', y_s'') and the background (x_b'', y_b''). The constant a is 0.101 for the group of sources that give red or yellow response to the observer (red-yellow recognition) and 0.051 for the group of sources that give green or blue response (green-blue recognition) for 50% detection in the range of background luminance 1,000−10,000 cd/m² corresponding to the luminance in the daytime. E_t is the achromatic threshold value of red light for red-yellow recognition or of green light for green-blue recognition against an achromatic background. In the same range of background luminance, the achromatic threshold value in lm/km² for 50% detection is given by

$$\log_{10} E_t = 0.9 \log_{10} B_b - b \tag{16}$$

where B_b is the background luminance expressed in cd/m², and b is 1.30 for red light and 1.07 for green light. The relations E_t *vs.* B_b and p *vs.* d are shown in Fig. 6.

In open air observation, natural objects make the background; mean values of

8) F. C. Breckenridge and W. R. Schaub: J. Opt. Soc. Amer. 29 (1939) 370.

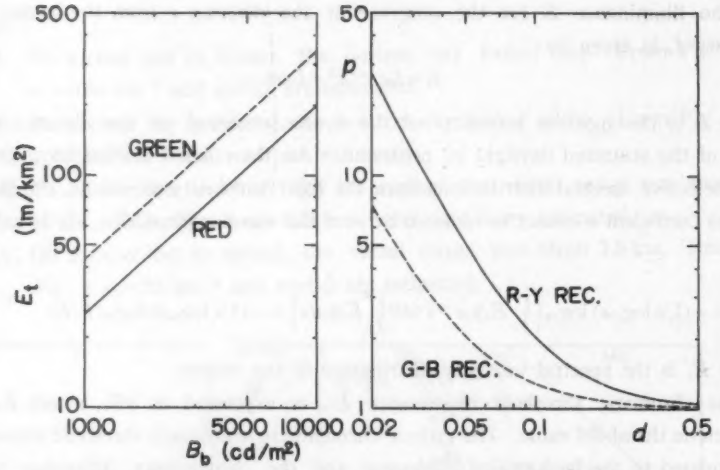


Fig. 6. Achromatic threshold illuminances of red and green point sources, E_t, as the funtion of background luminance, B_b, and ratios of chromatic to achromatic thresholds for red-yellow and green-blue recognitions against background of luminance 1000—10000 cd/m², p, as the function of chromaticity difference between source and background, d.

Table 2. Mean values of apparent luminance of natural objects (measured data).

Object	Distance (km)	Season	Weather	Luminance (cd/m²)
foliage	0–1	spring	clear	530
			overcast	240
		summer	clear	710
			overcast	380
		autumn	clear	550
			overcast	300
		winter	clear	730
			overcast	550
	6–8	spring	clear	1800
			overcast	450
	16	spring	clear	2050
grass	0.1	summer	overcast	880
		winter	clear	1580
field crops	0.1	summer	clear	1340
			overcast	370
earth	0.1	winter	clear	1100
			overcast	510
sea surface	1	winter	clear	1840
			overcast	880
snow surface	0.1	winter	clear	12400
building	2	spring	overcast	470
horizon sky		winter	clear	7100
		spring	overcast	2000

their luminance are shown in Table 2. The average values of their chromaticity coordinates reported in the previous paper are transformed to the RUCS coordinates and shown in Table 2.

Equating E given by (12) to E_{te} or pE_t, we obtain the visual range r_v from the equation

$$E_t/I_0=e^{-\sigma r_v}/pr_v^2 \tag{17}$$

For any combination of the source and background, $e^{-\sigma r}/pr^2$ is computed by (13), (14) and (15) and the $e^{-\sigma r}/pr^2$ *vs.* r curve is plotted on a diagram. Given the values of B_b and I_0, E_t/I_0 can be computed from (16) and the visual range can be obtained from the curve.

As examples, four sources, red (S 1), yellow (S 2), green (S 3) and pale blue (S 4) are used. Their spectral intensity distributions are shown in Fig. 8. As the state of the atmosphere three cases are given : case I′ ($\sigma_0=0.02\,\mathrm{km}^{-1}$, $n=3$), case II′ ($\sigma_0=0.20\,\mathrm{km}^{-1}$, $n=2$) and case III′ ($\sigma_0=2.0\,\mathrm{km}^{-1}$, $n=1$). The backgrounds are

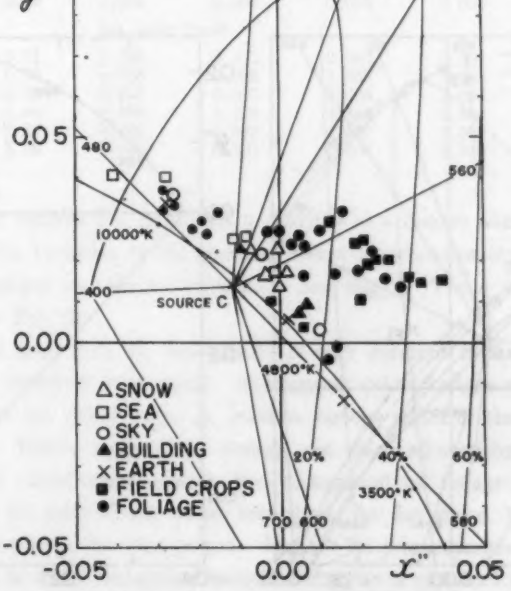


Fig. 7. Mean values of chromaticities of natural objects (RUCS coordinates).

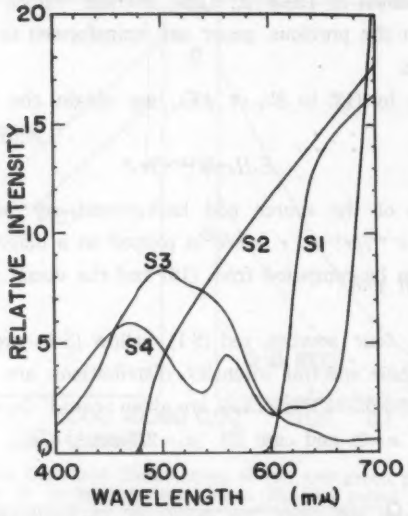


Fig. 8. Spectral intensity distribution curves of red (S 1) yellow (S 2), green (S 3), and pale blue (S 4) lights.

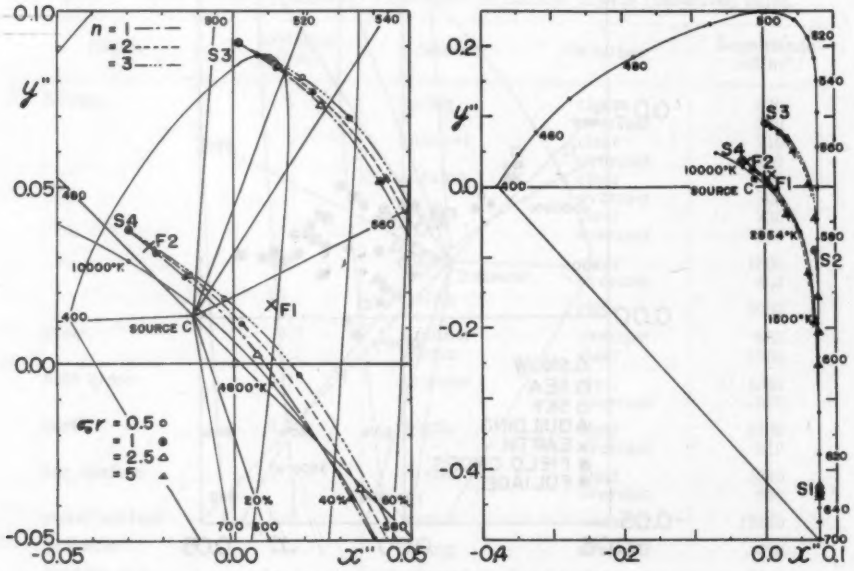


Fig. 9. RUCS coordinates of lights attenuated in the atmosphere (calculated results).

Table 3. Ratio of the apparent intensity of light sources observed at distance r to the initial intensity, I_r/I_0, (calculated results).

$\sigma_0 r$ \ n	0	1	2	3
	S 1 (red)			
0.01	0.990	0.992	0.993	0.993
0.10	0.905	0.919	0.928	0.939
0.50	0.607	0.650	0.689	0.724
1.00	0.369	0.422	0.474	0.524
2.50	0.082	0.116	0.154	0.200
5.00	0.007	0.013	0.024	0.042
	S 2 (yellow)			
0.01	0.990	0.991	0.992	0.992
0.10	0.905	0.909	0.912	0.915
0.50	0.609	0.622	0.635	0.645
1.00	0.368	0.385	0.401	0.416
2.50	0.083	0.093	0.105	0.118
5.00	0.007	0.009	0.011	0.016
	S 3 (green)			
0.01	0.991	0.991	0.990	0.990
0.10	0.905	0.904	0.902	0.900
0.50	0.608	0.602	0.596	0.590
1.00	0.368	0.363	0.357	0.350
2.50	0.082	0.080	0.079	0.078
5.00	0.006	0.006	0.006	0.006
	S 4 (pale blue)			
0.01	0.992	0.992	0.992	0.991
0.10	0.906	0.906	0.905	0.904
0.50	0.607	0.607	0.606	0.604
1.00	0.369	0.369	0.369	0.368
2.50	0.082	0.083	0.087	0.090
5.00	0.007	0.007	0.009	0.010

foliage in the vicinity F 1 (x=0.329, y=0.369) and a distant object F 2 (x=0.282, y=0.301). The variation of the apparent intensity and chromaticity of the sources in the atmosphere are shown in Table 3 and Fig. 9. The $e^{-\sigma r}/pr^2$ *vs.* r curves are shown in Fig. 10.

With red light (S 1), its chromaticity is very different from that of the background and therefore p is small. Its apparent color changes very little by the extinction and the visual range is good in various states of the atmosphere and background. Yellow light (S 2), though not much affected by the atmosphere, has a smaller visual range against the background of foliage or earth in comparison with red light of the same intensity. As for green light, it is fairly distinct in color from the background. Though its color changes by the selective extinction, it is easily recognized because the locus of chromaticity deviates fairly from natural objects. Green light of saturation lower than S 3 have poor visibility, for they become assimilated in color with the background. Pale blue light (S 4)

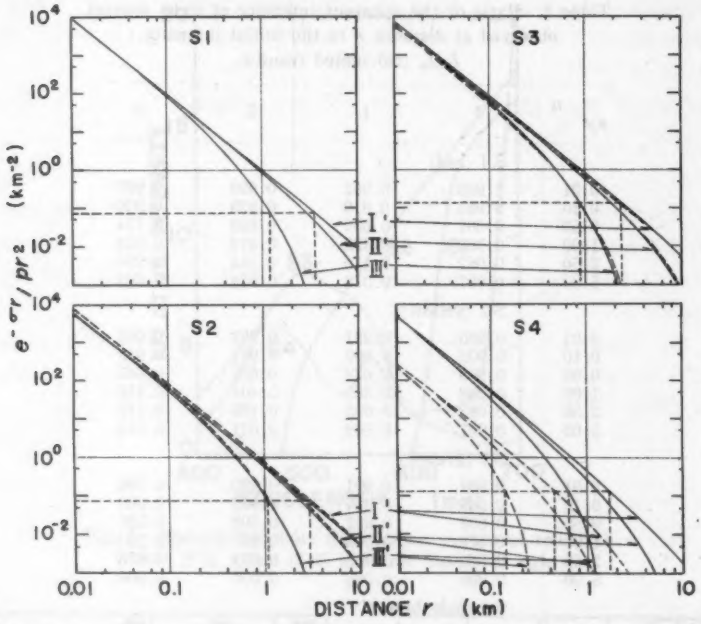


Fig. 10. $e^{-\sigma r}/pr^2$ as the function of distance r. Full lines represent observations against background F 1 and dashed lines observation against F 2.

has the least visual range, for it resembles the color of natural objects; it changes to white and yellow by selective extinction.

Given B_0=2,500 cd/m^2 and I_0=1,000 cd, E_t/I_0=0.077 for red-yellow recognition and 0.133 for green-blue recognition. Against these two backgrounds, F 1 and F 2, the visual ranges obtained from Fig. 10 are 1.2–3.4 km, 1.1–2.8 km and 0.9–2.1 km for S 1, S 2 and S 3 respectively. For S 4 the ranges are 0.6–1.7 km against F 1 and only 0.13–0.42 km against F 2.

5. Discussion

In the estimation of extinction coefficients from the change of colors, it is assumed that the spectral reflectances of measured objects are equal and the atmosphere is uniform. If equally finished targets are located at various distances, the first assumption is justified unless they fade differently. But such arrangement on a large scale would deprive the instrumental simplicity of its advantage. In the examples in section 3, foliage at various distances was actually measured. To investigate the error involved in this method, the change of colors is computed on two sets of initial values,

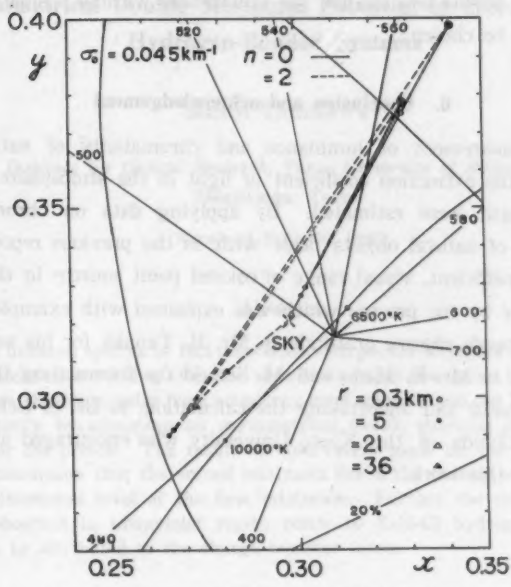


Fig. 11. Chromaticities of the light from foliage at various distances showing the effect of different initial values (calculated results).

B_{0i}, with the same B_{hi} and σ_0. The difference between the initial values of these two sets is thought to be maximum for foliage of the same kind measured at the same time. The results are shown in Fig. 11. The difference is fairly large at a small distance but decreases rapidly with the increase in distance. It is found that the colors of distant objects are affected by the color of horizon sky and the characteristic of extinction coefficients, especially their dependence on wavelength rather than by initial colors. As the computation of the visual range requires only rough values of σ_λ, it suffices to select objects of the same kind at various distances.

In the computation by (9), the spectral energy distributions of light from the horizon sky, B_{hi}, and from object near the observer, B_{0i}, are necessary. However, because the spectral intensity distributions of natural objects do not differ with one another greatly as shown in the previous paper, spectral distributions so corrected by the already known distributions for the object of the same kind as to be consistent with the measured values of the luminance and chromaticity can be applied. It causes a very small error. Therefore extinction coefficients are estimated by the measurement of apparent luminance and chromaticity alone.

To justify the second assumption, the atmosphere without visible layers of local cloud or fog must be chosen.

6. Conclusion and acknowledgement

From the measurement of luminance and chromaticity of natural objects at various distanc s, the extinction coefficient of light in the atmosphere and its dependence on wavelength were estimated. By applying data on chromatic threshold illuminance, colors of natural objects dealt with in the previous report and the estimated extinction coefficient, visual range of colored point sources in the daytime was obtained graphically by the process which was explained with examples.

The author extends sincere gratitude to Mr. H. Tanaka for his assistence in the experimental work, to Mr. K. Mano and M. Suzuki for formulating the program for the electronic computor and undertaking the calculation, to Dr. S. Uchida for his aid, and to Prof. Y. Uchida of the Kyoto University who encouraged and directed the author to complete this work.

Asymmetrical Double Minimum Potential in HCl-Solvents Hydrogen-Bonded Systems

Sanzoo TAKESAWA

Institute for Optical Research, Tokyo University of Education
Shinjiku-ku, Tokyo

(Received June 20, 1960)

Abstract

The infrared spectra of HCl associated with proton accepters in the region of fundamental stretching vibration reveal two broad absorption bands. These bands are shown to arise from an energy level splitting and can be interpreted satisfactorily by assuming an asymmetrical double potential in the energy curve for the proton. The result of observation made on the spectra leads to the conclusion that the second minimum lies in the vicinity of unperturbed (v=1) vibrational level of the first minimum. Further the new absorption bands observed in ultraviolet region relate to X-H-Cl hydrogen bond and seem to be attributed to the charge transfer force.

1. Introduction

As regards the existence of a second minimum in potential energy curve for a proton in a weak hydrogen bonded system, many theoretical[1)-8)] and experimental[9)-10)] studies have been carried out with much interest. Sato[7)] observed two new bands of the first OH stretching overtone in hydrogen-bonded solvents and gave a detailed interpretation by assuming a double minimum potential for the motion of proton. Recently Barrow and his coworkers[9) 10)] observed a similar splitting of the first OH stretching overtone of alcohols and concluded theoretically that a second potential minimum lies in the region of v=2 OH vibrational level. However, it is rather a

1) C. A. Coulson and V. Danielson, Arkiv. Fysik. **8** (1954) 239.
2) E. R. Lippincott and R. Schroeder, J. Chem. Phys. **23** (1955) 1099.
3) E. R. Lippincott, J. Chem. Phys. **26** (1957) 1678.
4) C. Reid, J. Chem. Phys. **30** (1959) 182.
5) P. C. Mckinery and G. M. Barrow, J. Chem. Phys. **31** (1959) 294.
6) Y. Sato and S. Nagakura, Science of Light. **4** (1955) 120.
7) Y. Sato, Science of Light. **6** (1957) 108.
8) R. Blinc and D. Hadzi, Mol. Phys. **1** (1958) 391.
9) C. L. Bell and G. M. Barrow, J. Chem. Phys. **31** (1959) 300.
10) C. L. Bell and G. M. Barrow, J. Chem. Phys. **31** (1959) 1158.

question whether $v=0$ or $v=1$ is the region in which the second potential minimum occurs. Although thorough examination of the OH fundamental absorption band would clear up the matter, correct measurment in this region is difficult because of the strong intensity of CH stretching absorption bands in solvents, while the separation of this absorption has been observed in OD fundamental band of phenols which is free from such interfering absorption.

In order to avoid the overlapping of CH stretching absorption in the fundamental region and further to enlighten us on the question of the height of the second potential minimum in X-H-Cl system, which is different from O-H-X system dealt with in the former studies, an experiment was made on the HCl fundamental absorption band in proton accepters. The remarkable shift of HCl vibrational absorption band in polar solvents has already been observed and interpreted as being attributed to hydrogen bonding by Martin, Gordy and Szobel,[11)-13)] but its splitting in infrared region of the spectrum has not been reported. Taylor and Vidale[14)] found two absorption bands in the Raman spectrum of 1:1 complex of HCl and dimethylether, but did not give appropriate interpretation. While measuring HCl fundamental absorption band of a few proton accepters on their infrared spectra, the auther found in every case a weak second absorption band on the long wavelength side of the strikingly shifted main band. For these two bands, a very satisfactory explanation was possible by assuming an asymmetrical double potential in X-H-Cl hydrogen bond formed between molecules of HCl and the solvent. The second minimum was found to lie $1800\,cm^{-1}$~$1900\,cm^{-1}$ higher than the vibrationless ($v=0$) level and lower than $v=1$ level of HCl and capable of interacting with the latter level. Between the height of this minimum and the relative basicity of proton accepters, a characteristic relation was found.

In describing the hydrogen bonding, the ionic and covalent nature of bonding electons and the role played by non-bonding electrons as well as the motion of protons should obviously be taken into consideration. On the absorption spectum of the complex of HCl and solvent in ultraviolet region, a new absorption band, which is not observed either in HCl or in the solvent, was found. The charge transfer force found out by Mulliken,[15)-17)] or in other words, the transition of non-bonding electrons of the proton accepter to an excited state level of HCl, is considered responsible for this new absorption band.

11) W. Gordy and P.C. Martin, J. Chem. Phys. **7** (1939) 99.
12) W. Gordy and P.C. Martin, J. Chem. Phys. **9** (1941) 215.
13) L. Szobel, Comp. rend. **218** (1944) 347, 843.
14) G.L. Vidale and R.C. Taylor, J. Am. Chem. Soc. **78** (1956) 294.
15) R.S. Mulliken, J. Chem. Phys. **7** (1939) 20.
16) R.S. Mulliken, J. Am. Chem. Soc. **72** (1950) 600.
17) R.S. Mulliken, J. Am. Chem. Soc. **74** (1952) 811.

2. Experimental and results

For obtaining absorption spectra in infrared region, a Koken D-301 spectrometer with two NaCl prism and 25~50 μ rocksalt cell, and in ultraviolet region, a Hitachi EPU-2 spectrometer with a qurtz prism and qurtz cell with 1 cm path length were used. In infrared region, frequency calibration was made for each spectrum by superposing polystyrene and carbon monoxide. HCl was used as the proton donor and several compounds such as ethers and ketones with lone pair electrons as the proton accepters. Molar concentration of HCl in solvents was made reasonably low for 1:1 complex to be formed. Care was taken of impurities, particularly water for it is likely to form hydrogen bond with the solvent.

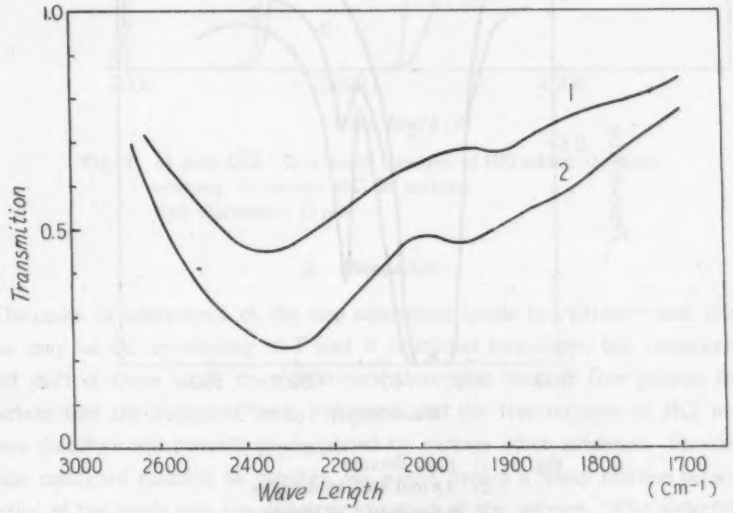


Fig. 1. 1) 1.9 mol HCl in diethylether
2) 3.8 mol HCl in dioxane
Cell thickness; 25 μ

Table

	ν_a	ν_b	$\Delta\nu$
Diethylether	2340 cm^{-1}	1895 cm^{-1}	455 cm^{-1}
Isopropylether	2370	1966	404
Dioxane	2322	1945	377
Acetone	2319	1969	351
Tetrahydrofuran	2231	1891	331

On infrared absorption spectra, two broad bands, one strong in the region of 2400 cm^{-1} and the other weak in the region of 1900 cm^{-1}, were always observable. The absorption curves of HCl-dioxane and HCl-diethylether solution are illustrated in Fig. 1 as examples, and the wavelength, separation and relative intensity of the two bands appeared in other cases are given in Table with curves in Fig. 4.

Further, on dioxane and diethylether, to which HCl was added, a sharp band was observed at 837 cm^{-1} and 826 cm^{-1} respectively. The band in the case of dioxane-HCl solution is shown in Fig. 2.

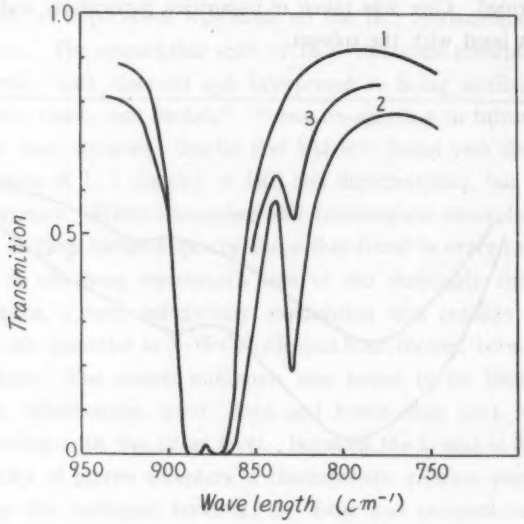


Fig. 2. 1) pure dioxane
2) 3.8 mol HCl in dioxane
3) 0.8 mol HCl in dioxane
Cell thickness: 25 μ

In ultraviolet absorption spectra of proton accepters with HCl, a broad absorption band was found in the neighborhood of 360 mμ, except for diethylether and isopropylether of which the maxima were obscure. The maximum of this absorption band was at 365 mμ for methylethylketone, at a little above 360 mμ for acetone, at 360 mμ for ethylacetate, at a little below 360 mμ for dioxane. On the other hand in inactive solvent carbon tetrachloride, the absorption band was at the wavelength shorter than 300 mμ overlapping the characteristic absorption band of carbon tetrachloride as clearly shown in Fig. 3.

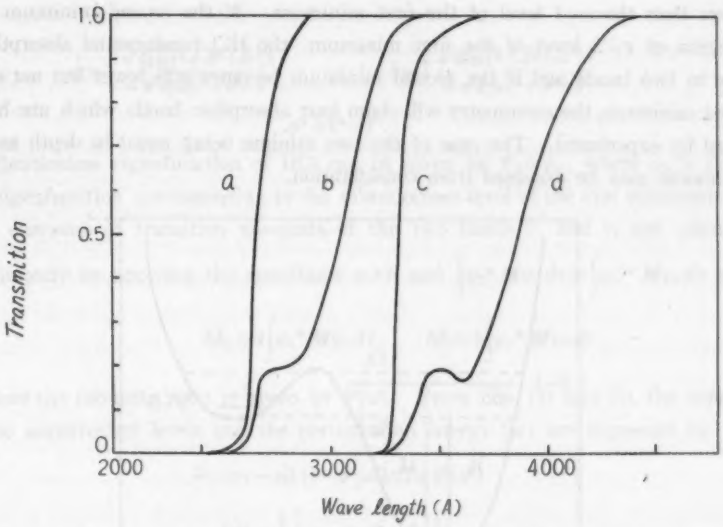


Fig. 3. a) pure CCl_4 b) a small amount of HCl added c) pure acetone d) 1.6 mol HCl in acetone
Cell thickness; 10 mm

3. Discussion

The cause of appearance of the two absorption bands in 2400 cm^{-1} and 1900 cm^{-1} regions may be the broadening of P and R rotational structures, but considering the marked shift of these bands from 2890 cm^{-1} absorption band of free gasious HCl, we are certain that the hydrogen bond is formed, and the free rotation of HCl molecule becomes therefore not possible as supported by various other evidence. Should however the restricted rotation be possible, we would expect a linear relation between the separation of the bands and the dielectric constant of the solvent. The dielectric constant of acetone is 21.3 which is very large in comparison with 2.2 of dioxane, yet the separations of the bands in these are of the same order which denies the said proportionality and hence, no such rotation.

On the other hand, the asymmetrical double minimum offers a very good explanation for the appearance of the band pair. When a double minimum in X-H-Cl hydrogen-bonded system is assumed, the concept that the fundamental absorption band of HCl has two components, is that the unperturbed ground level ($v=0$) of the second minimum is at a position where it can interact with the unperturbed $v=1$ level of the first minimum. The fact that the intensity of the absorption band on the long wavelength side is weaker makes us infer that the $v'=0$ level of the second minimum

is lower than the $v=1$ level of the first minimum. If the second minimum lies in the region of $v=2$ level of the first minimum, the HCl fundamental absorption can not be in two bands and if the second minimum becomes still lower but not equal to the first minimum, the asymmetry will claim four absorption bands which are however negated by experiment. The case of the two minima being equal in depth as we see on ammonia may be excepted from consideration.

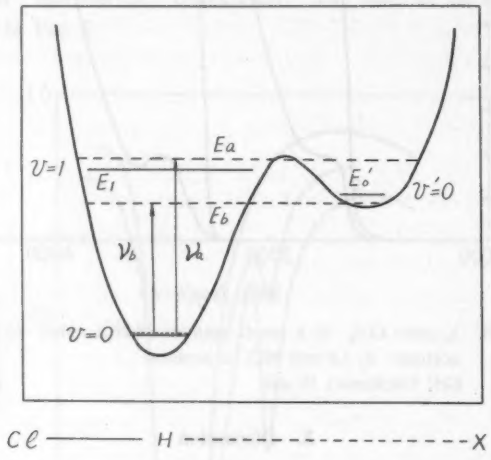


Fig. 4.

Fig. 4 illustrates schematically a double minimum potential; the level is on arbitary scale. As a result of perturbation one obtains for the perturbed energy level E_a and E_b[18]

$$E_a=\frac{1}{2}(E_1+E_0')+\sqrt{4|w|^2+\delta^2} \tag{1a}$$

$$E_b=\frac{1}{2}(E_1+E_0')-\sqrt{4|w|^2+\delta^2} \tag{1b}$$

where E_1 and E_0' are unperturbed energy level values and $\delta=E_1-E_0'$ is the separation of the unperturbed levels. w is the exchange integral $\int\varphi_1^* H\varphi_0 d\tau$. The eigenfunctions of the two resulting levels E_a and E_b are mixtures of zero order eingenfunctions φ_1 and φ_0' corresponding to E_1 and E_0' respectively. They are

$$\Psi_a=a\varphi_1-b\varphi_0' \tag{2a}$$

$$\Psi_b=a\varphi_0'+b\varphi_1 \tag{2b}$$

18) G. Herzberg, Spectra of Diatomic Molecules, P283 (D. Van. Nostrand, New York).

where

$$a=\left\{\frac{\sqrt{4|w|^2+\delta^2}+\delta}{2\sqrt{4|w|^2+\delta^2}}\right\}^{\frac{1}{2}} \qquad b=\left\{\frac{\sqrt{4|w|^2+\delta^2}+\delta}{2\sqrt{4|w|^2+\delta^2}}\right\}^{\frac{1}{2}} \tag{3}$$

and

$$a^2+b^2=1\,. \tag{4}$$

The vibrationless eigenfunction of HCl can be given by $\Psi_0\simeq\varphi_0$, where φ_0 is the zero order eigenfunction corresponding to the vibrationless level of the first minimum. The matrix elements of transition moments of the two bands ν_a and ν_b are calculated[7] approximately by applying the conditions $a>b$ and $\int\varphi_1^*\boldsymbol{M}\varphi_0 d\tau\gg\int\varphi_0'^*\boldsymbol{M}\varphi_0 d\tau$ as

$$M_a\doteqdot a\int\varphi_1^*\boldsymbol{M}\varphi_0 d\tau\,, \qquad M_b\doteqdot b\int\varphi_1^*\boldsymbol{M}\varphi_0 d\tau\,. \tag{5}$$

Therefore the intensity ratio is given by b^2/a^2. From Eqs. (1) and (3), the separation δ of the unperturbed levels and the perturbation energy $|w|$ are expressed by

$$\delta=(\nu_a-\nu_b)\,(1-b^2/a^2)/(1+b^2/a^2) \tag{6}$$

$$|w|=\frac{1}{2}\left\{(\nu_a-\nu_b)^2-\delta^2\right\}^{\frac{1}{2}}. \tag{7}$$

The relations between the basicity constant[19] of proton accepters and values of δ and $|w|$ calculated by using the values of b^2/a^2 and $\Delta\nu=\nu_a-\nu_b$ are illustrated in Fig. 5. Along with the decrease of proton attractive force of solvent, the values of δ show

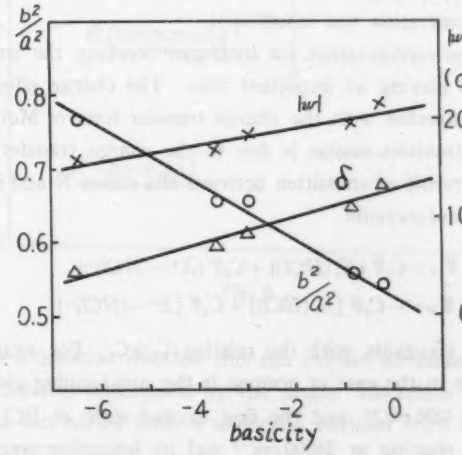


Fig. 5. The relations between the basicity constant and values of δ, $|w|$ and b^2/a^2.

19) W. Gordy and S. C. Stanford, J. Chem. Phys. **9** (1941) 204.

a tendency to fall off as seen in Fig. 5 suggesting the rise of level of the second minimum relative to the first and with it the decrease in difference between a and b as can be understood from Eq. (3). This means more φ_1 and φ_0' to mix, but the hydrogen bond becomes weaker.

The experiment confirms the increase of intensity of the band on the long wavelength side due to the intensity ratio b^2/a^2 being very much dependent on δ rather than $|w|$ as the basicity of solvent is reduced. $|w|$ relates to the height of the potential barrier between the two minima. Although this height is diffcult to be estimated, it is reasonable to conclude that such a barrier exists and such systems as are dealt with here do not show a single potential minimum and that the height of the second minimum is in the vicinity of 1800 cm^{-1}~1900 cm^{-1} from the vibrationless ($v=0$) level of the first minimum.

Beside the already discussed broad absorption band, a sharp band appeared in the cases of diethylether and dioxane at 826 cm^{-1} and 837 cm^{-1} respectively by the addition of HCl. This band increases in intensity along with HCl molar concentration. It has been known that the skeletal vibrational frequency shifts when molecular complex is formed.[20] The above mentioned new band is probably attributed to a low frequency shift of symmetrical C–O–C skeletal frequency in the hydrogen bond. Adams and Katz[21] found on dioxane a similar band at about 840 cm^{-1} by the addition of HF. This proves also the formation of O–H–Cl hydrogen bond. With substances other than diethylether and dioxane, the frequency shift was not conspicuous probably because HCl molar concentration was insufficient.

As to the electron configuration for hydrogen bonding, the structure X^+–H : Cl^- should be regarded as playing an important role. The charge effect of non-bonding pair of electrons is connected with the charge transfer force of Mulliken. If the new absorption band in ultraviolet region is due to the charge transfer force, we should consider it to be the result of transition between the states N and E corresponding to the following two eigenfunctions

$$\Psi_N = C_1\Psi\{X, (HCl)\} + C_2\Psi\{X^+ - (HCl)^-\} \tag{8a}$$

$$\Psi_E = -C_2\Psi\{X, (HCl)\} + C_1\Psi\{X^+ - (HCl)^-\} \tag{8b}$$

where C_1 and C_2 are constants with the relation $C_1 \gg C_2$. For example, the highest electron-occupied state in the case of acetone is the non-bonding electron state and its ionization potential is 9.96 eV,[22] and the first excited state of HCl molecule appears as a continuous band starting at 44000 cm^{-1} and its ionization potential is 12.7 eV.[22]

20) A. D. E. Pullin and J. Mc. Pollock, Trans. Faraday Soc. **54** (1958) 11.
21) R. M. Adams and J. J. Katz, J. Mol. Spectr. **1** (1957) 309.
22) K. Watanabe, J. Chem, Phys. **26** (1957) 542.

The exact position of the maximum of this continuous band has not been make known, but by analogy with the case of HBr band, we may be allowed to assign it above 50000 cm^{-1}. The values of ionization potential given above are for acetone and HCl both in free gas phase; in solutions they may be different but their relative positions are as illustrated in Fig. 6. The $n \rightarrow \pi^*$ absorption band resulting from C=O of acetone (2790 A in hexane solution)[23][24] shifts because of the hydrogen bonding. But this band is different from the $\pi \rightarrow \pi^*$ band[25][26] and in most case does not show the "red shift" when the hydrogen bond is formed. Therefore the appearence of the band at about 360 mμ owes likely to the charge transfer in the molecular complex.

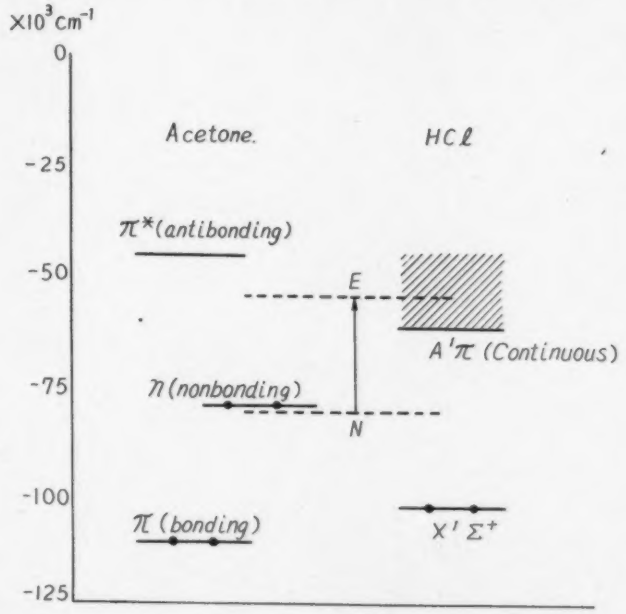


Fig. 6.

However, it is a question whether this can explain adequately the case of absorption in CCl_4–HCl solution superposed by the proper absorption of CCl_4. CCl_4 being an inactive solvent can hardly form a molecular complex with HCl molecule and the

23) H. L. McMurry and R. S. Mulliken, Pro. Nat. Acad. Sci. **26** (1940) 312.
24) H. L. McMurry, J. Chem. Phys. **9** (1941) 231, 241.
25) S. E. Sheppard, Rev. Mod. Phys. **14** (1942) 303.
26) H. McConnel, J. Chem. Phys. **20** (1952) 700.

band in question may in fact be the band proper to CCl_4 or HCl but shifted. Further, it is strange that the absorption bands of diethylether and isopropylether-HCl solution have no maximum. For the present, as far as the absorption band that appears near 360 mμ in the cases of acetone-HCl etc. is concerned, the charge transfer may be responsisible for the absorption provided that the band is not caused by impurities.

Aknowledgment

The auther is sincerely grateful to Prof. H. Ootuka and Prof. K. Kudo for their encouragement and interest, to Dr. Y. Sato for his valuable suggestion and advice.

CONTENTS